

OM nucleic - nucleic search, using sw model

```
run on: January 19, 2003, 19:50:40 ; Search time 2218 Seconds
        (without alignments)
        9324.453 Million cell updates/sec
```

```
Title: US-09-844-864-1
Perfect score: 127
Sequence: 1 aaagcgcgacgagccgcggaa.....acaaaaaaaaaaaaaaa 127
```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```

post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database :

|     |                |
|-----|----------------|
| 1:  | em_estha:*     |
| 2:  | em_esthnm:*    |
| 3:  | em_estln:*     |
| 4:  | em_estlun:*    |
| 5:  | em_estlov:*    |
| 6:  | em_estlpi:*    |
| 7:  | em_estro:*     |
| 8:  | em_hnc:*       |
| 9:  | gb_estl:*      |
| 10: | gb_estl2:*     |
| 11: | gb_hnc:*       |
| 12: | gb_est3:*      |
| 13: | gb_est4:*      |
| 14: | gb_est5:*      |
| 15: | em_estfun:*    |
| 16: | em_estlom:*    |
| 17: | gb_gss:*       |
| 18: | em_gss_hnv:*   |
| 19: | em_gss_hum:*   |
| 20: | em_gss_p1n:*   |
| 21: | em_gss_vrl:*   |
| 22: | em_gss_fun:*   |
| 23: | em_gss_mam:*   |
| 24: | em_gss_mus:*   |
| 25: | em_gss_other:* |
| 26: | em_gss_pro:*   |
| 27: | em_gss_rod:*   |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

## SUMMARIES

| Result | No. | Score | Query | Match | length | DB       | ID       | Description |
|--------|-----|-------|-------|-------|--------|----------|----------|-------------|
| C      | 1   | 646.2 | 50.6  | 673   | 12     | BC071693 | BC071693 | H3102B03    |
|        |     | 643   | 50.4  | 675   | 10     | BB641267 | BB641267 | H3102B03    |
|        | 2   | 642.4 | 50.3  | 666   | 10     | BB626422 | BB626422 | H3102B03    |
|        | 3   | 580.4 | 44.5  | 593   | 12     | BC084538 | BC084538 | H3102B03    |
|        | 4   | 570.4 | 44.7  | 594   | 12     | BE471866 | BE471866 | UI-M-BH3    |
|        | 5   | 497.2 | 38.9  | 525   | 10     | BB704019 | BB704019 | BB704019    |
|        | 6   |       |       |       |        |          |          |             |

|   |    |       |      |     |    |           |           |           |
|---|----|-------|------|-----|----|-----------|-----------|-----------|
| C | 7  | 473.8 | 37.1 | 499 | 9  | A1854700  | A1854700  | UT-M-BHO- |
| C | 8  | 448.4 | 35.1 | 491 | 10 | BB703259  | BB703259  | BB703259  |
| C | 9  | 413.8 | 32.4 | 491 | 10 | BB704648  | BB704648  | BB704648  |
| C | 10 | 405.8 | 31.8 | 436 | 9  | A0044294  | A0044294  | A0044294  |
| C | 11 | 398.2 | 31.2 | 425 | 10 | BB705931  | BB705931  | BB705931  |
| C | 12 | 396.4 | 31.0 | 590 | 9  | A0023153  | A0023153  | A0023153  |
| C | 13 | 393   | 30.8 | 450 | 12 | BE946858  | BE946858  | UT-M-BH3  |
| C | 14 | 392.2 | 30.7 | 419 | 10 | BB704449  | BB704449  | BB704449  |
| C | 15 | 367.4 | 28.8 | 521 | 10 | BB703869  | BB703869  | BB703869  |
| C | 16 | 367   | 28.7 | 427 | 10 | BB706957  | BB706957  | BB706957  |
| C | 17 | 364.2 | 28.5 | 434 | 10 | BB699732  | BB699732  | BB699732  |
| C | 18 | 347.2 | 27.2 | 391 | 10 | BB704146  | BB704146  | BB704146  |
| C | 19 | 333.8 | 26.1 | 440 | 10 | BB700620  | BB700620  | BB700620  |
| C | 20 | 316.2 | 24.8 | 441 | 10 | BB704848  | BB704848  | BB704848  |
| C | 21 | 264.6 | 20.7 | 315 | 10 | BB248342  | BB248342  | BB248342  |
| C | 22 | 224.8 | 17.6 | 654 | 17 | A2225053  | A2225053  | RPCI-23-6 |
| C | 23 | 221.6 | 16.6 | 251 | 10 | AV3358045 | AV3358045 | AV3358045 |
| C | 24 | 208.6 | 16.3 | 251 | 10 | AV337766  | AV337766  | AV337766  |
| C | 25 | 203.2 | 15.9 | 249 | 10 | BB063113  | BB063113  | BB063113  |
| C | 26 | 203   | 15.9 | 290 | 10 | BB060381  | BB060381  | BB060381  |
| C | 27 | 200   | 15.7 | 286 | 10 | BB060134  | BB060134  | BB060134  |
| C | 28 | 199.8 | 15.6 | 246 | 10 | BM861479  | BM861479  | BM861479  |
| C | 29 | 197.8 | 15.5 | 642 | 14 | BM98465   | BM98465   | BM98465   |
| C | 30 | 197   | 15.4 | 643 | 13 | BM035466  | BM035466  | BM035466  |
| C | 31 | 194.6 | 15.2 | 484 | 13 | BB666714  | BB666714  | BB666714  |
| C | 32 | 194.6 | 15.2 | 532 | 13 | BB667744  | BB667744  | BB667744  |
| C | 33 | 194.6 | 15.2 | 536 | 13 | BB676515  | BB676515  | BB676515  |
| C | 34 | 194.6 | 15.2 | 571 | 13 | BB676757  | BB676757  | BB676757  |
| C | 35 | 194.6 | 15.2 | 603 | 13 | BB676161  | BB676161  | BB676161  |
| C | 36 | 194.6 | 15.2 | 617 | 13 | BB666863  | BB666863  | BB666863  |
| C | 37 | 194.6 | 15.2 | 624 | 13 | BB666863  | BB666863  | BB666863  |
| C | 38 | 194.2 | 15.2 | 373 | 9  | A1217763  | A1217763  | A1217763  |
| C | 39 | 192   | 15.0 | 247 | 10 | AV358133  | AV358133  | AV358133  |
| C | 40 | 191.2 | 15.0 | 542 | 13 | BB680099  | BB680099  | BB680099  |
| C | 41 | 186.6 | 14.6 | 229 | 10 | AV359620  | AV359620  | AV359620  |
| C | 42 | 184.2 | 14.4 | 512 | 12 | BG729074  | BG729074  | BG729074  |
| C | 43 | 184.2 | 14.4 | 542 | 13 | BB43179   | BB43179   | BB43179   |
| C | 44 | 181   | 14.0 | 209 | 13 | BB061815  | BB061815  | BB061815  |
| C | 45 | 179   | 14.0 | 179 | 12 | BB663658  | BB663658  | BB663658  |

## ALIGNMENTS

28071693/c  
 LOCUS BG071693 673 bp mRNA  
 DEFINITION AB0102803-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
 ACCESSION AB0102803.3 mRNA sequence.  
 VERSION BG071693.1 GI:12554262  
 KEYWORDS none

**SOURCE**

ORGANISM

## REFERENCE

AUTHORS

# TITLE

COMMENT

TABLE 1.10.2

—

1



|                       |                 |                     |           |             |
|-----------------------|-----------------|---------------------|-----------|-------------|
| Query Match           | 50.4%;          | Score 643;          | DB 10;    | Length 675; |
| Best Local Similarity | 97.6%;          | Pred. No. 2.5e-102; |           |             |
| Matches 660;          | Conservative 0; | Mismatches 15;      | Indels 1; | Gaps 1;     |

|    |     |        |        |        |        |       |       |       |       |     |      |        |      |
|----|-----|--------|--------|--------|--------|-------|-------|-------|-------|-----|------|--------|------|
| Db | 601 | ACCAGT | NAGTG  | -TACTT | CAACAN | GTCTG | CGCAG | TGTGT | GAGAA | ATC | NTAC | ACCTTA | 6599 |
| QY | 946 | CAGACT | GCAGAC | ATTC   | 961    |       |       |       |       |     |      |        |      |
|    |     |        |        |        |        |       |       |       |       |     |      |        |      |
| Db | 660 | CAGAGT | GCAGAC | ATTC   | 675    |       |       |       |       |     |      |        |      |

| RESULT 3 | LOCUS    | DEFINITION  | ACCESSTION |
|----------|----------|---|------------|
| BB264222 | BB264222 | 666 bp mRNA<br>RIKEN full-length enriched, 10 days neonate cortex Mus | BB264222   |
|          |          | musculus CDNA clone AB30014H23 3', mRNA sequence.                     | AB264222   |

| KEYWORDS | EST.         |
|----------|--------------|
| SOURCE   | house mouse. |
| ORGANISM | Mus musculus |

REFERENCES  
AUTHORS

On Jul 7, 2000 this sequence version replaced v1.9050570

104  
AUTHOR  
TITLE  
JOURNAL  
COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>  
Carninci, P., Shibata, Y., Hayamatsu, N., Sugahara, Y., Shibata, K., Itoh,  
M., Kono, H., Okazaki, Y., Muramatsu, M., Hayashizaki, Y.  
Normalization and subtraction of cdp-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
Wael, K., Fujiwara, S., Inoue, K., Togawa, I., Izawa, M., Ohara, E.,  
Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.

RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11): 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shiota, K. and Iwawashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome, 12, 673-677 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

## FEATURES

```
1.000
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A830014H23"
```

|                       |       |              |          |            |
|-----------------------|-------|--------------|----------|------------|
| Query Match           | 50.3% | Score 642.4  | DB 10    | Length 666 |
| Best Local Similarity | 99.6% | Pred. No. 3  | 2e-102   |            |
| Matches 665           | 0     | Mismatches 1 | Indels 2 | Gaps 2     |

|                           |   |                                    |      |
|---------------------------|---|------------------------------------|------|
| QY                        | 1248  | GCTATTCG                           | 1255 |
| Db                        | 659   | GCTATTGC                           | 666  |
| RESULT 4                  |   |                                    |      |
| LOCUS                     | BG084538  |                                    |      |
| DEFINITION                | BG084538  | 593 bp mRNA linear EST 26-JAN-2001 |      |
| ACCESSION                 | H3102B03.5  | 5' mRNA sequence.                  |      |
| VERSION                   | BG084538  |                                    |      |
| KEYWORDS                  | BG084538.1  | GI:12567102                        |      |
| SOURCE                    | EST.  |                                    |      |
| ORGANISM                  | house mouse.  |                                    |      |
| REFERENCE                 | Mus musculus  |                                    |      |
| AUTHORS                   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;<br>Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  |                                    |      |
| TITLE                     | Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradot, S.A., Tanaka<br>, T.S., Carter, M.G. and Ko, M.S.H.<br><u>Verification and initial annotation of NIA mouse 15K cDNA clone set</u><br><u>published 12001</u>   |                                    |      |
| JOURNAL                   | <del>Unpublished</del>  |                                    |      |
| COMMENT                   | <del>Other: BGN H3102B03-3</del><br>Contact: George J. Kargul<br>Laboratory of Genetics<br>National Institute on Aging/National Institutes of Health<br>333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA<br>Email: cdna@sgsun.grc.nia.nih.gov<br>This clone set has been freely distributed to the community. Please<br>visit <a href="http://lgsun.grc.nia.nih.gov/cDNA/15K.html">http://lgsun.grc.nia.nih.gov/cDNA/15K.html</a> for details.<br>Plate: H3102 row: B column: 03<br>Seq primer: -21M13 Reverse<br>High quality sequence stop: 593<br>POLYA-No.  |                                    |      |
| FEATURES                  | Location/Qualifiers   |                                    |      |
| SOURCE                    | 1..593<br>/organism="Mus musculus"<br>/strain="C57BL/6J"<br>/db_xref="taeST:H3102B03-5"<br>/db_xref="taxon:10090"<br>/clone="H3102B03"<br>/clone_1lb="NIA Mouse 15K cDNA Clone Set"<br>/sex="Clones arrayed from a variety of cDNA libraries"<br>/dev_stage="Clones arrayed from a variety of cDNA<br>libraries"<br>/lab_host="DH10B"<br>/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI. This<br>clone is among a rearranged set of 15,247 clones from 11<br>embryo cDNA libraries (including preimplantation stage<br>embryos from unfertilized egg to blastocyst, embryonic<br>part of E7.5 embryos, extraembryonic part of E7.5 embryos<br>, and E12.5 female mesonephros/gonad) and one newborn<br>ovary cDNA library. Average insert size 1.5 kb. All<br>source libraries are cloned unidirectionally with Oligo(dT)<br>) Not primers. References include: (1) Genome-wide<br>expression profiling of mid-gestation placenta and embryo<br>using a 15,000 mouse developmental cDNA microarray,<br>Proc. Natl. Acad. Sci. U S A, 97: 9127-9133; (2)<br>Large-scale cDNA analysis reveals phased gene expression<br>patterns during preimplantation mouse development, 2000<br>Development, 127: 1737-1749; (3) Genome-wide mapping of<br>unselected transcripts from extraembryonic tissue of<br>7.5-day mouse embryos reveals enrichment in the t-complex<br>and under-representation on the X chromosome, 1998, Hum<br>Mol Genet 7: 1967-1978." |                                    |      |
| BASE COUNT                | 142 a 165 c 192 g 94 t  |                                    |      |
| ORIGIN                    |   |                                    |      |
| Query Match               | 45.5%; Score 580.4; DB 12; Length 593;  |                                    |      |
| Best Local Similarity     | 99.7%; Pred. No. 2e-91;   |                                    |      |
| Matches 592; Conservative | 0; Mismatches 1; Indels 1; Gaps 1   |                                    |      |

Db 1 CGGATCTCTGGGGAGACCGGGAACCGGAGCCGAGAGAGTGCCCGCGAGAAAGCGTCC 60  
 Oy 572 CCCAGCCCGGAGCGGAGGCGATGTTTCAGGCTGCGAGGAGCGCGGTGGAGACG 631  
 Db 61 CCCAGCCCGGAGCGGAGGCGATGTTTCAGGCTGCGAGGAGCGCGGTGGAGACG 120  
 Oy 632 ACCCAACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 691  
 Db 121 AGCCACACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
 Oy 692 AGCCATGTCCTGCGCCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 751  
 Db 181 AGCCATGTCCTGCGCCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
 Oy 752 AGCCATGTCCTGCGCCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 811  
 Db 241 AGCCATGTCCTGCGCCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
 Oy 812 AGAGCTACGCTACTATCATCTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 871  
 Db 301 AGAGCTACGCTACTATCATCTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 360  
 Oy 872 GGTGTCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 931  
 Db 361 GGTGTCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 419  
 Oy 932 TCCACACACCTTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 991  
 Db 420 TCCACACACCTTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 479  
 Oy 992 GCTGTCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1051  
 Db 480 GCTGTCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 539  
 Oy 1052 GCGAGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1105  
 Db 540 GCGAGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 593

RESULT 5  
 LOCUS BF471866 594 bp mRNA linear EST 04-DEC-2000  
 DEFINITION UI-M-BH3-bwu-b-08-0-UI.r1 NIH-BMAP\_M.S4 Mus musculus cDNA clone  
 ACCSSION BF471866  
 VERSION BF471866.1 GI:11541049  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 594)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 92044477  
 COMMENT Contact: Chln, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mestr@mail.nih.gov  
 CDNA Library Preparation: M.B. Soares Lab Clone distribution:  
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
 should be noted that Bento Soares is generating a small number of  
 additional specialized non-redundant arrays of BMAP cDNAs whose  
 availability will be considered under appropriate and limited  
 collaborative arrangements  
 Seq primer: M3 Reverse.  
 Location/Qualifiers

## FEATURES

Location/Qualifiers

## source

1. 594  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BH3-bwu-b-08-0-UI"  
 /clone\_1lb="NIH-BMAP\_M.S4"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified  
 polylinker. Site 1: Not I; Site 2: Eco RI. The  
 NIH-BMAP\_M.S4 library is a subtracted library of a series,  
 ultimately derived from a mixture of individually tagged  
 normalized libraries from ten regions of the mouse brain  
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,  
 cortex, amygdala, basal ganglia, pineal gland, striatum,  
 hippocampus) after a series of subtractions to reduce the  
 representation of cDNAs from which ESTs had already been  
 generated. The following serially subtracted libraries  
 were generated in this process: NIH-BMAP\_M.S4,  
 NIH-BMAP\_M.S3.3, NIH-BMAP\_M.S3.2, NIH-BMAP\_M.S3.1,  
 NIH-BMAP\_M.S2, NIH-BMAP\_M.S1. The subtracted library  
 (NIH-BMAP\_M.S4) was constructed as follows: PCR amplified  
 cDNA inserts from NIH-BMAP\_M.S3.3, NIH-BMAP\_M.S3.2, and  
 NIH-BMAP\_M.S3.1 clones from which 3' ESTs had been derived  
 was used as a driver in a hybridization with a pool of  
 the NIH-BMAP\_M.S3.3, NIH-BMAP\_M.S3.2, and NIH-BMAP\_M.S3.1  
 libraries in the form of single-stranded circles. The  
 remaining single-stranded circles (subtracted library)  
 was purified by hydroxyapatite column chromatography,  
 converted to double-stranded circles and electroporated  
 into DH10B bacteria (Life Technologies) to generate the  
 NIH-BMAP\_M.S4 library. This procedure has been previously  
 described (Bonaldo, Lennon and Soares, Genome Research  
 6:791-806, 1996)"

Paste Comp 105 a 184 g 158 g 147 t

Query Match 44.7% Score 570.4; DB 12; Length 594;  
 Best Local Similarity 99.5%; Pfd. No. 1.1e-89;  
 Matches 593; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Oy 555 CCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 614  
 Db 594 CCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 535  
 Oy 615 GCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 674  
 Db 534 GCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 475  
 Oy 675 TGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 734  
 Db 474 TGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 415  
 Oy 735 GGATGCCCCCTGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 794  
 Db 414 GGATGCCCCCTGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 355  
 Oy 795 TTTCAGTTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 854  
 Db 354 TTTCAGTTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 295  
 Oy 855 GGAGAGCCGCTATGTCGTGTGTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 914  
 Db 294 GGAGAGCCGCTATGTCGTGTGTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 236  
 Oy 915 GCCGAGTGTGAGAAATCTTACAAACCTTACAGAGTGGAGGAGGAGGAGGAGGAGG 974  
 Db 235 GCCGAGTGTGAGAAATCTTACAAACCTTACAGAGTGGAGGAGGAGGAGGAGGAGG 176  
 Oy 975 GTAAAGAACTAGATGTGTCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1034  
 Db 175 GTAAAGAACTAGATGTGTCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 116

|            |   |  |      |  |
|------------|---|--|------|--|
| Oy         | 1035  | ATCGGCAAGACTTGTGTGGGAGATGCAAGGACAAACGCCGTGCTCGAGCAGACCTTGA   | 1094 |  |
| Db         | 115   | ATCGGCAAGACTTGTGTGGGAGATGCAAGGACAAACGCCGTGCTCGAGCAGACCTTGA   | 56   |  |
| Oy         | 1095  | GCTTCAATATCATATTAGTAGAGTGCAGAAACGTTTCCTAGATGGGGCTAAT   | 1150 |  |
| Db         | 55  | GCTTCAATATCATATTAGTAGAGTGCAGAAAC - TTTCTGTAGATGGGGCTAAT  | 1    |  |
| RESULT 6   |   |  |      |  |
| LOCUS      | BB704019  | 525 bp   | mrna |  |
| DEFINITION | BB704019  | RIKEN full-length enriched, in vitro fertilized eggs Mus musculus cDNA clone 7420459B08 3', mRNA sequence. |      |  |
| ACCESSION  | BB704019  |  |      |  |
| VERSION    | BB704019.1  | GI:16052854  |      |  |
| KEYWORDS   | EST.  |  |      |  |
| SOURCE     | house mouse.  |  |      |  |
| ORGANISM   | Mus musculus.   |  |      |  |
| REFERENCE  |   |  |      |  |
| AUTHORS    | 1 (bases 1 to 525)  |  |      |  |
|            | Akimura,T., Arikawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayashizaki,Y., Hirooka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku,A., Akhita,S., Tanaka,T., Tomaru,A., Toya,T., Matakichi,A., Yasunishi,A., Muramatsu,M., and Hayashizaki,Y.   |  |      |  |
| TITLE      | RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)   |  |      |  |
| JOURNAL    | Unpublished (2001)  |  |      |  |
| COMMENT    | Contact: Yoshihide Hayashizaki<br>Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute<br>The Institute of Physical and Chemical Research (RIKEN)<br>1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan<br>Tel.: 81-45-503-9222<br>Fax: 81-45-503-9216<br>Email: genome-res@gsc.riken.go.jp/<br>URL: http://genome-gsc.riken.go.jp/<br>Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.<br>, Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)<br>waji,K., Fujiwake,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E., Maitiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.<br>RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)<br>Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y., and Hayashizaki,Y.<br>Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)<br>Please visit our web site (http://genome-gsc.riken.go.jp) for further details.<br>e mouse tissues.<br>Location/Qualifiers<br>1. 525<br>/organism="Mus musculus"<br>/strain="C57BL/6J"<br>/db_xref="taxon:10090"<br>/clone_lib="7420459B08"<br>/clone_lib="RIKEN full-length enriched, in vitro fertilized eggs"<br>/sex="female"<br>/tissue_type="in vitro fertilized eggs"<br>/dev_stage="egg" |  |      |  |

```

/lab.host="DH108"
/NOTE="Site_1: Sall; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer 15'
GACAGAGAGAGATCCAGAGACTCTTTTGTGTGTGTGTATGTA 3'. cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence 15'
GAGGAGAGATCTCTCAGATTATTAATTAATCCCCCCCCCC 3'. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified Bluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI"
BASE COUNT      142 a      134 c      129 g      120 t
ORIGIN
Query Match      38.9%; Score 497.2; DB 10; Length 525;
Best Local Similarity 99.0%; Pred. No. 6,1e-77;
Matches 521; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY      732  TTTCGATGCCCCCTCGAGACCAAGCCCTCCCGCAAAAGCAGGAGCAGACAGAGCCCT 791
          TTTTCGATGCCCCCTCGAGACCAAGCCCTCCCGCAAAAGCAGGAGCAGACAGAGCCCT 791
Db      1  TTTCGATGCCCCCTCGAGACCAAGCCCTCCCGCAAAAGCAGGAGCAGACAGAGCCCT 60
          TTTTCGATGCCCCCTCGAGACCAAGCCCTCCCGCAAAAGCAGGAGCAGACAGAGCCCT 60

QY      792  GCGTTTCCAGTTCTTAAGAGCAGAACTACGGGTACTATACATGCAGAGACTCAAAATTCG 851
          GCGTTTCCAGTTCTTAAGAGCAGAACTACGGGTACTATACATGCAGAGACTCAAAATTCG 851
Db      61  GCGTTTCCAGTTCTTAAGAGCAGAACTACGGGTACTATACATGCAGAGACTCAAAATTCG 120
          GCGTTTCCAGTTCTTAAGAGCAGAACTACGGGTACTATACATGCAGAGACTCAAAATTCG 120

QY      852  GTGGAGAGACGCCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 911
          GTGGAGAGACGCCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 911
Db      121  GTGGAGAGACGCCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 179
          GTGGAGAGACGCCCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 179

QY      912  TCTGCCGAGTGTGTGAGAAATCTCAACCCCTTAAGAGTGGAGACATCTCAACCTGTCAA 971
          TCTGCCGAGTGTGTGAGAAATCTCAACCCCTTAAGAGTGGAGACATCTCAACCTGTCAA 971
Db      180  TCTGCCGAGTGTGTGAGAAATCTCAACCCCTTAAGAGTGGAGACATCTCAACCTGTCAA 239
          TCTGCCGAGTGTGTGAGAAATCTCAACCCCTTAAGAGTGGAGACATCTCAACCTGTCAA 239

QY      972  GTTTGTTAAAGAACTAGATGTGCTTCCCAAGTCAGATTTTGCACAGCTGGACCTTAAAGCC 1031
          GTTTGTTAAAGAACTAGATGTGCTTCCCAAGTCAGATTTTGCACAGCTGGACCTTAAAGCC 1031
Db      240  GTTTGTTAAAGAACTAGATGTGCTTCCCAAGTCAGATTTTGCACAGCTGGACCTTAAAGCC 299
          GTTTGTTAAAGAACTAGATGTGCTTCCCAAGTCAGATTTTGCACAGCTGGACCTTAAAGCC 299

QY      1032  CCCATCGCAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1091
          CCCATCGCAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1091
Db      300  CCCATCGCAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 359
          CCCATCGCAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 359

QY      1092  TCAGCTTCAAAATCAATCATTTAGTGTGAGATGCAGAAACGTTTGTGTAGATGGGCTAATG 1151
          TCAGCTTCAAAATCAATCATTTAGTGTGAGATGCAGAAACGTTTGTGTAGATGGGCTAATG 1151
Db      360  TCAGCTTCAAAATCAATCATTTAGTGTGAGATGCAGAAACGTTTGTGTAGATGGGCTAATG 419
          TCAGCTTCAAAATCAATCATTTAGTGTGAGATGCAGAAACGTTTGTGTAGATGGGCTAATG 419

QY      1152  GAATGACAGAACTGAGCTTTCTCCCTCTTCCCTTCCCTTCCCTTCCCAATTTCTTCATGACA 1211
          GAATGACAGAACTGAGCTTTCTCCCTCTTCCCTTCCCTTCCCTTCCCTTCCCAATTTCTTCATGACA 1211
Db      420  GAATGACAGAACTGAGCTTTCTCCCTCTTCCCTTCCCTTCCCTTCCCAATTTCTTCATGACA 479
          GAATGACAGAACTGAGCTTTCTCCCTCTTCCCTTCCCTTCCCTTCCCAATTTCTTCATGACA 479

QY      1212  GACAGTGTACTTGATATATAAGCCGTGAATATAAAGGATTGGCAA 1257
          GACAGTGTACTTGATATATAAGCCGTGAATATAAAGGATTGGCAA 1257
Db      480  GACAGTGT-TACTTGGATATAAAGCCGTGAATATAAAGGATTGGCAA 524
          GACAGTGT-TACTTGGATATAAAGCCGTGAATATAAAGGATTGGCAA 524

RESULT 7
A1854700/c  A1854700      499 bp      mRNA      linear      EST 15-JUL-1999
LOCUS      UI-M-BHO-akc-d-12-0-UI.s1 NIH_BMAP_M.S1 Mus musculus cDNA clone
DEFINITION UI-M-BHO-akc-d-12-0-UI 3', mRNA sequence.
ACCESSION A1854700
VERSION    A1854700.1 GI:5498606
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

```

REFERENCE 1 (bases 1 to 499)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT

Contact: Chih, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7150, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: MSSTEMail.nih.gov  
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized basal ganglia library cDNA library preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.  
 Seq primer: M13 Forward  
 PolyA-yes

## FEATURES

## SOURCE

Location/Qualifiers  
 1. 499  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="U1-M-BH0-alc-d-12-0-U1"  
 /clone\_lib="NIH\_BMAP\_M\_S1"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT7T3D-Pac (Pharmacia)"  
 /note="Site 1: Not I; Site 2: Eco RI; The NIH\_BMAP\_M\_S1 library is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 20,000 cDNA clones obtained from non-normalized and normalized libraries of these ten regions of the mouse brain.  
 TAG-LIB=NIH\_BMAP\_M\_S1  
 TAG-TISSUE=basal-ganglia  
 TAG\_SEQ=TGTAC"  
 BASE COUNT 115 a 120 c 117 g 147 t

Query Match 37.1% Score 473.8; DB 9; Length 499;  
 Best Local Similarity 99.2% Pred. No. 7.2e-73;  
 Matches 497; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
 Oy 772 GAGCAGCAAGAGGCGCGCTTCCACTTCTAGACGAGAATGCGCTACTATCAC 831  
 Db 499 GAGCAGCAAGAGGCGCGCTTCCACTTCTAGAGGAGAATGCGCTACTATCAC 440  
 Oy 832 TGCAGAGCTGCAGAAATCCGGTGGAGAGCGCTATGTGTGTGAGGAGCACCAGT 891  
 Db 439 TGCAGAGCTGCAGAAATCCGGTGGAGAGCGCTATGTGTGTGAGGAGCACCAGT 380  
 Oy 892 AAGCTTACTTTCAGACAGTTTCCGAGAGTGTGAGAAATCTTACACCTTACAGAGT 951  
 Db 379 AAGCTT-TACTTTCAGACAGTTTCCGAGAGTGTGAGAAATCTTACACCTTACAGAGT 321  
 Oy 952 GAGAGCATCATCCTGTCAAGTTGTAAAGACTAGATGCTGCCAGTGCAGATTTCG 1011  
 Db 320 GAGAGCATCATCCTGTCAAGTTGTAAAGACTAGATGCTGCCAGTGCAGATTTCG 261  
 Oy 1012 CCACGTGAGACCTTAAAGCCGCCATGGCAAGACTTGTGTGGAGATGCAAGACAAAG 1071  
 Db 260 CCACGTGAGACCTTAAAGCCGCCATGGCAAGACTTGTGTGGAGATGCAAGACAAAG 201

Oy 1072 CCGTCTCGGACAGCAGCCTTTCAGCTTCAATATCATCATTTAGTGAAGTCGAACGTT 1131  
 Db 200 CCGTCTCGGACAGCAGCCTTTCAGCTTCAATATCATCATTTAGTGAAGTCGAACGTT 141  
 Oy 1132 TCTGTAGATGGGCTTAATGGAATGCAAGTAGAGCTTTCCTCCCTTCACCTTCC 1191  
 Db 140 TCTGTAGATGGGCTTAATGGAATGCAAGTAGAGCTTTCCTCCCTTCACCTTCC 81  
 Oy 1192 TTTCCAAATTCCTTCATGACAGACAGTGTACTTGATATTAAGCCCTGTAAATAGTA 1251  
 Db 80 TTTCCAAATTCCTTCATGACAGACAGTGTACTTGATATTAAGCCCTGTAAATAGTA 22  
 Oy 1252 TTGCAAAACAAAAA 1272  
 Db 21 TTGCAAAACAAAAA 1

RESULT 8  
 BB703259 491 bp mRNA linear EST 11-OCT-2001  
 BB703259 RIKEN full-length enriched, in vitro fertilized eggs Mus  
 musculus cDNA clone 7420449J15.3', mRNA sequence.  
 BB703259  
 VERSION BB703259.1 GI:16052094  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 491)  
 AUTHORS Akimura,T., Arikawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imctani,K., Ishii,Y., Ito,M., Kawai,D., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Nunasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takeku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanabe,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)

JOURNAL  
 COMMENT Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Science Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,  
 URL: http://genome-gsc.riken.go.jp/  
 Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)

Waghi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,K., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara Y. and Hayashizaki,Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome-gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.

FEATURES  
 Source Location/Qualifiers  
 1. 491







**CONTACT:** HIROJUMI DOI  
 DOI Biosymmetry Project, ERATO  
 Japan Science and Technology Corporation (JST)  
 WBC Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan  
 Email: [hdq@doi.jst.go.jp](mailto:hdq@doi.jst.go.jp)  
 Location/Qualifiers

|           |  |
|-----------|--|
| KEYWORDS  | EST.   |
| SOURCE    | house mouse.   |
| ORGANISM  | Mus musculus   |
| REFERENCE | Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 425) |

**AUTHORS**  
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
Hayatsu, N., Hiramoto, K., Hiraka, T., Hirozane, T., Imotani, K., Ishii



RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

Query Match 31.0%; Score 396.4; DB 9; Length 590;  
Best Local Similarity 86.1%; Pred. No. 1.9e-59;  
Matches 464; Conservative 0; Mismatches 71; Indels 4; Gaps 3;

BASE COUNT 135 a 133 c 156 g 149 t 17 others  
ORIGIN

QY 713 TGGCTCAGACACCCGGGATTCGGATGCCCTCGAGACCGCCCTCCCAAGCAGCG 772  
DB 542 TGGCTCAGACACCCGGGATTCGGATGCCCTCGAGACCGCCCTCCCAAGCAGCG 483  
QY 773 AGCGAGCAG 890  
DB 482 AGNCAAG 423  
QY 831 CTGCAAG 889  
DB 422 CTGCAAG 363  
QY 890 GTAGAGTCTACTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 949  
DB 362 AGTAAAGTCTACTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 303  
QY 950 GTGAGAGACATCATCTGCTCAAAAGTGTAAAGAGAGAGAGAGAGAGAGAGAGAG 1009  
DB 302 GTGAGAGACATCATCTGCTCAAAAGTGTAAAGAGAGAGAGAGAGAGAGAGAGAG 243  
QY 1010 CGCCAGCTGAGACCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1069  
DB 242 CGCCAGCTGAGACCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183  
QY 1070 CGCCTGCTGAG 1129  
DB 182 CGCCTGCTGAG 123  
QY 1130 TTTCTGCTAGATGGGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1189  
DB 122 TTTCTGCTAGATGGGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63  
QY 1190 CTTTCCCAAAATTTCTTCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248  
DB 62 CTTTCCCAAAATTTCTTCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5

RESULT 13  
BE946858 450 bp mRNA linear EST 03-OCT-2000  
LOCUS  
DEFINITION  
ACCESSION BE946858  
VERSION BE946858.1 GI:10524617  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

Genome Res. 6 (9), 791-806 (1996)  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mstetmell.nih.gov  
Oligo-dT track not found. Not I site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.B.  
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA  
clones from RESEARCH GENERICS. It should be noted that Bento Soares  
is generating a small number of additional specialized  
non-redundant arrays of BMAP cDNAs whose availability will be  
considered under appropriate and limited collaborative arrangements  
Seq primer: M13 Forward  
POLYA=No.

FEATURES  
Source

Location/Qualifiers  
1. 450  
/organism="Mus musculus"  
/strain="C57BL/6j"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH3-awu-b-08-0-UI"  
/clone\_lib="NIH\_BMAP\_M.S4"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: p773D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not I; Site 2: Eco RI. The  
NIH\_BMAP\_M.S4 library is a subtracted library of a series,  
ultimately derived from a mixture of individually tagged  
normalized libraries from ten regions of the mouse brain  
(cerebellum, brain stems, olfactory bulbs, hypothalamus,  
cortex, amygdala, basal ganglia, pineal gland, striatum,  
hippocampus) after a series of subtractions to reduce the  
representation of cDNAs from which ESTs had already been  
generated. The following serially subtracted libraries  
were generated in this process: NIH\_BMAP\_M.S4,  
NIH\_BMAP\_M.S3.3, NIH\_BMAP\_M.S3.2, NIH\_BMAP\_M.S3.1,  
NIH\_BMAP\_M.S2, NIH\_BMAP\_M.S1. The subtracted library  
(NIH\_BMAP\_M.S4) was constructed as follows: PCR amplified  
cDNA inserts from NIH\_BMAP\_M.S3.3, NIH\_BMAP\_M.S3.2, and  
NIH\_BMAP\_M.S3.1 clones from which 3' ESTs had been derived  
was used as a driver in a hybridization with a pool of  
the NIH\_BMAP\_M.S3.3, NIH\_BMAP\_M.S3.2, and NIH\_BMAP\_M.S3.1  
libraries in the form of single-stranded circles. The  
remaining single-stranded circles (subtracted library)  
was purified by hydroxyapatite column chromatography,  
converted to double-stranded circles and electroporated  
into DH10B bacteria (Life Technologies) to generate the  
NIH\_BMAP\_M.S4 library. This procedure has been previously  
described (Bonaldo, Lennon and Soares, Genome Research  
6:791-806, 1996)  
TAG\_SEQ=None found"

BASE COUNT 62 a 168 c 156 g 63 t 1 others  
ORIGIN

Query Match 30.8%; Score 393; DB 12; Length 450;  
Best Local Similarity 99.7%; Pred. No. 8.2e-59;  
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCGGCGAGGCGGCGGAGACAGACAGATCTCCCGGAGACAGCTTCACCCCTGCCGCA 62  
DB 57 GCGCGGCGAGGCGGCGGAGACAGACAGATCTCCCGGAGACAGCTTCACCCCTGCCGCA 116  
QY 63 TCCCTATCCGAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122  
DB 117 TCCCTATCCGAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176  
QY 123 ACCCGGCGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182  
DB 177 ACCCGGCGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 236  
QY 183 CAGCCAGCAGGCGGAG 242  
DB 237 CAGCCAGCAGGCGGAG 296  
QY 243 CAGCCAGCAGGCGGAG 302  
DB 297 CAGCCAGCAGGCGGAG 356  
QY 303 CCGGCGGCGGAG 362

Db 357 CGGGGGCGCAGCGTGCAGCGCTGCAGGGGCGCGACCGCCGAGCCCGCATCGGTTCC 416

Oy 363 CTGTCAACCCCGTGGCCACAGCCGGCGCGGAGAGA 396  
|||||  
|||||

Db 417 CTGTCAACCCCGTGGCAGCGCGCGGAGAGA 450

RESULT 14

LOCUS BR704449 419 bp mRNA linear EST 11-OCT-2001

DEFINITION BR704449 RIKEN full-length enriched, in vitro fertilized eggs Mus musculus cDNA clone 7420464A04 3', mRNA sequence.

ACCESSION BR704449

VERSION BR704449.1 GI:16053284

KEYWORDS EST.

ORGANISM house mouse.  
Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,H., Toyata,R., Watanishi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y., RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)

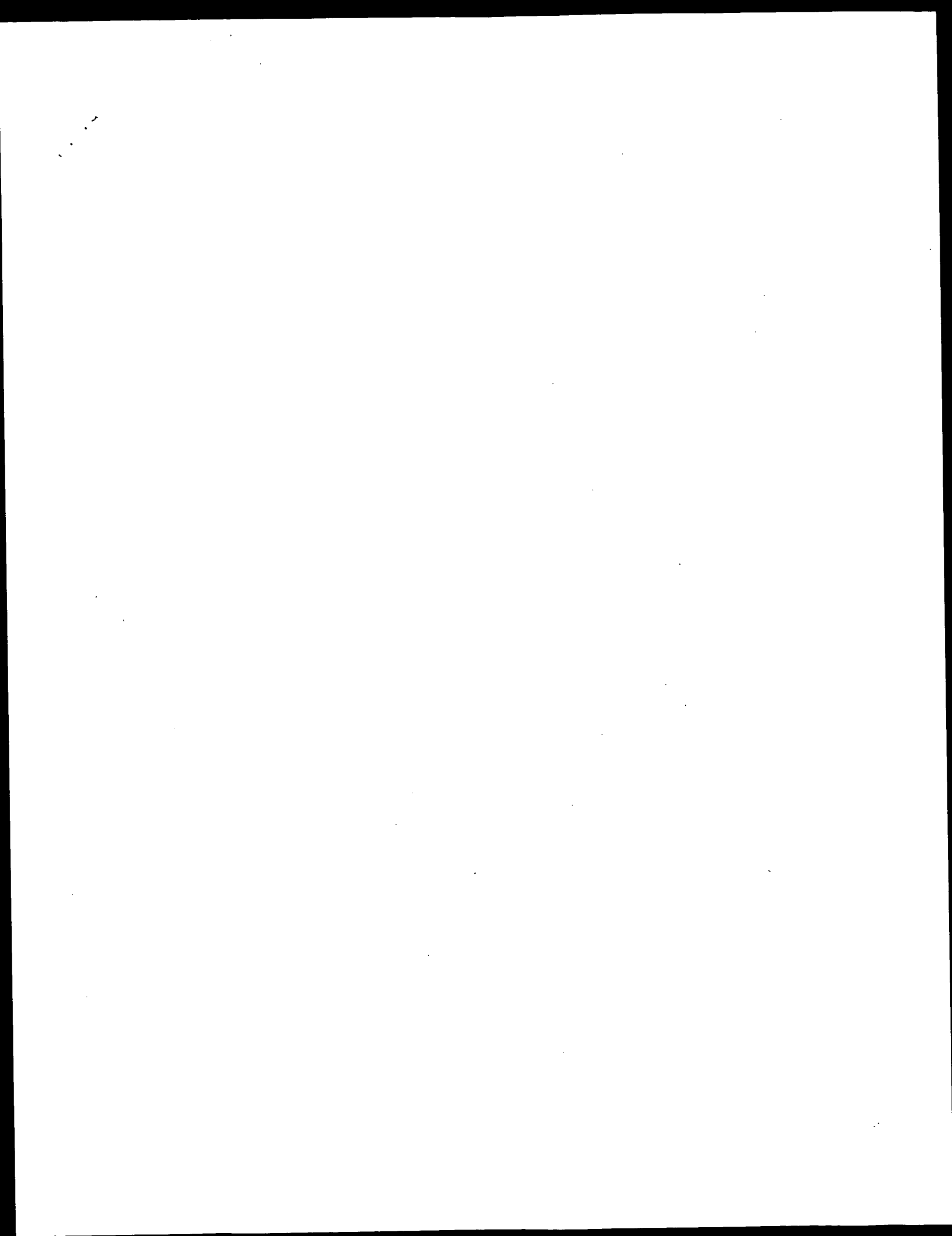
TITLE Unpublished (2001)

JOURNAL Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y., Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagii,K., Fujiwaka,S., Inoue,K., Togawa,Y., Iawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y., RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y., Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.  
e mouse tissues.

FEATURES  
Source location/Qualifiers  
1. 419  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="7420464A04"  
/clone\_lib="RIKEN full-length enriched, in vitro fertilized eggs"  
/sex="female"  
/tissue\_type="in vitro fertilized eggs"  
/dev\_stage="egg"  
/lab\_host="DH10B"  
/note="Site\_1: SalI, Site\_2: BamHI; cDNA library was

|                       |   |   |           |          |         |      |              |    |            |    |      |   |        |    |
|-----------------------|---|---|-----------|----------|---------|------|--------------|----|------------|----|------|---|--------|----|
|                       | BEST LOCAL SIMILARITY   | 98.8%   | PRED. NO. | 1-2E-58: | MATCHES | 416; | CONSERVATIVE | 0; | MISMATCHES | 3; | GAPS | 2;  | INDELS | 2; |
| BASE COUNT            | 115 A   | 100 C   | 100 G     | 104 T    |         |      |              |    |            |    |      | FPLC I. CLONING SITES, 5' END: SALL; 3' END: BANHI* |        |    |
| ORIGIN                |   |   |           |          |         |      |              |    |            |    |      |   |        |    |
| Query Match           | 30.7% ; Score 392.2; DB 10; Length 419;   |   |           |          |         |      |              |    |            |    |      |   |        |    |
| Best Local Similarity | 98.8%; Pred. No. 1.2e-58;   |   |           |          |         |      |              |    |            |    |      |   |        |    |
| Matches               | 416; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  |   |           |          |         |      |              |    |            |    |      |   |        |    |
| OY                    | 834   | CAAGACATGCATAAAATCCGGTGGGAGACGCCCTATTGTGTGTGTGCAGGGCACCATTA     | 893       |          |         |      |              |    |            |    |      |   |        |    |
| Dd                    | 1   | CAAGACTCGCAAATTCGCGTGGAAGAGCGCTATGTGTGTGTGTGCAGGGCACCATTA       | 60        |          |         |      |              |    |            |    |      |   |        |    |
| OY                    | 894   | GGTTGTACTTCAACAAGTTTTGCGAGTGTGTGAGAANAATCTCAACACCCTTAGAGATGG    | 953       |          |         |      |              |    |            |    |      |   |        |    |
| Dd                    | 61  | GGT-TTACCTTCAAACAGTTTCTGCCCCAAGTTTGTGAAAAATCTCAACACCCTTAGAGATGG | 119 .     |          |         |      |              |    |            |    |      |   |        |    |
| OY                    | 954   | AGGACATACCGTCCAAGTTGTAAAACATCAGATGTGCCGCCCACATCAAGATTTCGCC      | 1013      |          |         |      |              |    |            |    |      |   |        |    |
| Dd                    | 120   | AGGACATACCGTCCAAGTTGTAAAACATCAGATGTGTGCCGCCCACATCAAGATTTCGCC    | 179       |          |         |      |              |    |            |    |      |   |        |    |
| OY                    | 1014  | ACGTGGACCCTAAACGCCGCCCATCGGCAAGAAGCTGTGTGGGAGATGCAAGAACAAAGCC   | 1073      |          |         |      |              |    |            |    |      |   |        |    |
| Dd                    | 180   | ACGTGGACCCTAAACGCCGCCCATCGGCAAGAAGCTGTGTGGGAGATGCAAGAACAAAGCC   | 239       |          |         |      |              |    |            |    |      |   |        |    |
| OY                    | 1074  | TGTCTCGGACAGACACCTTCACGCTTCAAAATCATCATTTAGTGAAGTGCAAAACGTTTC    | 1133      |          |         |      |              |    |            |    |      |   |        |    |
| Dd                    | 240   | TGTCTCGGACAGACACCTTCACGCTTCAAAATCATCATTTAGTGAAGTGCAAAACGTTTC    | 299       |          |         |      |              |    |            |    |      |   |        |    |
| OY                    | 1134  | TGCTGATGGGGGCTAATGAATGGACAAGTAGAGCTTCTCCGCTCTTACCTCTTCCCTT      | 1193      |          |         |      |              |    |            |    |      |   |        |    |
| Dd                    | 300   | TGCTGATGGGGGCTAATGAATGGACAAGTAGAGCTTCTCCGCTCTTACCTCTTCCCTT      | 359       |          |         |      |              |    |            |    |      |   |        |    |
| OY                    | 1194  | TCCAAATCTTCATGACAGACAGTGTACTTGATATATAAGCCTGTGAATAAAGTATT        | 1253      |          |         |      |              |    |            |    |      |   |        |    |
| Dd                    | 360   | TCCAAATCTTCATGACAGACAGTGTACTTGATATATAAGCCTGTGAATAAAGTATT        | 418       |          |         |      |              |    |            |    |      |   |        |    |
| OY                    | 1254  | C 1254  |           |          |         |      |              |    |            |    |      |   |        |    |
| Dd                    | 419   | G 419   |           |          |         |      |              |    |            |    |      |   |        |    |
| RESULT 15             |   |   |           |          |         |      |              |    |            |    |      |   |        |    |
| LOCUS                 | BB703869  | 521 bp mRNA linear EST 11-Oct-2001                              |           |          |         |      |              |    |            |    |      |   |        |    |
| DEFINITION            | BB703869 RIKEN full-length enriched, in vitro fertilized eggs Mus musculus cDNA clone 7420457c21 3', mRNA sequence.   |   |           |          |         |      |              |    |            |    |      |   |        |    |
| ACCESSION             | BB703869  |   |           |          |         |      |              |    |            |    |      |   |        |    |
| VERSION               | BB703869.1 GI:16052704  |   |           |          |         |      |              |    |            |    |      |   |        |    |
| KEYWORDS              | EST.  |   |           |          |         |      |              |    |            |    |      |   |        |    |
| SOURCE                | house mouse,  |   |           |          |         |      |              |    |            |    |      |   |        |    |
| ORGANISM              | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Mus.   |   |           |          |         |      |              |    |            |    |      |   |        |    |
| REFERENCE             | Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,R., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi.K., Nomura,K., Numasaki.R., Okazaki.Y., Okido.T., Salito,R., Sakai,C., Sakai,K., Sakazume.N., Sasaki.D., Satou,K., |   |           |          |         |      |              |    |            |    |      |   |        |    |





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: January 19, 2003, 20:01:55 ; Search time 3366 Seconds

(Without alignments)  
11041.082 Million cell updates/sec

|                |                 |
|----------------|-----------------|
| Title:         | US-09-844-864-1 |
| Perfect score: | 1277            |

**Sequence:**

1 aaagcgcgcgacgcgcgcga.....acaaaaaa 1277

Scoring table: IDENTITY\_NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 08

Listing first 45 summaries

Database

|     |                 |
|-----|-----------------|
| 1:  | gb_ba:*         |
| 2:  | gb_htg:*        |
| 3:  | gb_in:*         |
| 4:  | gb_ov:*         |
| 5:  | gb_ov:*         |
| 6:  | gb_pat:*        |
| 7:  | gb_ph:*         |
| 8:  | gb_pl:*         |
| 9:  | gb_pr:*         |
| 10: | gb_ro:*         |
| 11: | gb_sbs:*        |
| 12: | gb_sy:*         |
| 13: | gb_un:*         |
| 14: | gb_vl:*         |
| 15: | em_ba:*         |
| 16: | em_fun:*        |
| 17: | em_hum:*        |
| 18: | em_in:*         |
| 19: | em_mu:*         |
| 20: | em_mus:*        |
| 21: | em_or:*         |
| 22: | em_ov:*         |
| 23: | em_pat:*        |
| 24: | em_ph:*         |
| 25: | em_pl:*         |
| 26: | em_ro:*         |
| 27: | em_scs:*        |
| 28: | em_un:*         |
| 29: | em_vl:*         |
| 30: | em_bhg_hum:*    |
| 31: | em_bhg_in:*     |
| 32: | em_bhg_other:*  |
| 33: | em_bhg_mus:*    |
| 34: | em_bhg_pln:*    |
| 35: | em_bhg_rod:*    |
| 36: | em_bhg_mam:*    |
| 37: | em_bhg_vrt:*    |
| 38: | em_hy:*         |
| 39: | em_hygo_hum:*   |
| 40: | em_hygo_mus:*   |
| 41: | em_hygo_other:* |

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query | Length | DB | ID        | Description         |
|------------|-------|-------|--------|----|-----------|---------------------|
| 1          | 729.4 | 62.6  | 200535 | 10 | AL671880  | AL671880 Mouse DNA  |
| 2          | 689.8 | 54.70 | 144963 | 2  | AC107666  | AC107666 Mus muscu  |
| 3          | 657.4 | 48.15 | 67535  | 2  | AC122733  | AC122733 Mus muscu  |
| 4          | 615.4 | 48.2  | 173007 | 2  | AC125993  | AC125993 Rattus no  |
| 5          | 585.2 | 45.8  | 89551  | 2  | AC127083  | AC127083 Rattus no  |
| 6          | 452.4 | 35.4  | 169908 | 2  | AC126519  | AC126519 Rattus no  |
| 7          | 360   | 28.2  | 212848 | 2  | AC108848  | AC108848 Mus muscu  |
| 8          | 289.8 | 23.5  | 212848 | 2  | AC108848  | AC108848 Mus muscu  |
| 9          | 194.8 | 15.3  | 89551  | 2  | AC127083  | AC127083 Rattus no  |
| 10         | 147.6 | 11.6  | 138890 | 2  | AC007970  | AC007970 Homo sapi  |
| 11         | 147.6 | 11.6  | 151961 | 2  | AC037443  | AC037443 Homo sapi  |
| 12         | 147.6 | 11.6  | 204268 | 2  | AC092163  | AC092163 Homo sapi  |
| 13         | 134.6 | 10.5  | 187578 | 2  | AC096552  | AC096552 Homo sapi  |
| 14         | 73.2  | 5.7   | 694    | 9  | HS3135296 | HS3135296 Homo sapi |
| 15         | 72.2  | 5.7   | 125020 | 9  | AF429315  | AF429315 Homo sapi  |
| 16         | 66.8  | 5.2   | 146431 | 2  | AC109670  | AC109670 Homo sapi  |
| 17         | 64.4  | 5.0   | 172433 | 2  | AC118973  | AC118973 Homo sapi  |
| 18         | 63.2  | 4.9   | 125020 | 9  | AF429315  | AF429315 Homo sapi  |
| 19         | 62.8  | 4.9   | 194873 | 2  | AC125563  | AC125563 Rattus no  |
| 20         | 62.0  | 4.9   | 1393   | 11 | PM111126  | PM111126 Homo sapi  |
| 21         | 62.0  | 4.9   | 136551 | 2  | AC048354  | AC048354 Homo sapi  |
| 22         | 62    | 4.9   | 138467 | 2  | AC111697  | AC111697 Rattus no  |
| 23         | 59    | 4.6   | 18563  | 2  | AC129765  | AC129765 Rattus no  |
| 24         | 58.6  | 4.6   | 194936 | 2  | AC115680  | AC115680 Mus muscu  |
| 25         | 58.4  | 4.6   | 36394  | 1  | SCD35     | SCD35 Homo sapi     |
| 26         | 58    | 4.5   | 92310  | 2  | AC105380  | AC105380 Homo sapi  |
| 27         | 58    | 4.5   | 97065  | 2  | AC121735  | AC121735 Homo sapi  |
| 28         | 57.2  | 4.5   | 62847  | 2  | AC125560  | AC125560 Rattus no  |
| 29         | 57.2  | 4.5   | 83629  | 2  | AC101485  | AC101485 Rattus no  |
| 30         | 57.2  | 4.5   | 144053 | 2  | AC092857  | AC092857 Rattus no  |
| 31         | 57.2  | 4.5   | 150090 | 2  | AC121750  | AC121750 Rattus no  |
| 32         | 57.2  | 4.5   | 150338 | 2  | CNS086A9  | CNS086A9 Rattus no  |
| 33         | 57    | 4.5   | 92055  | 2  | AC105692  | AC105692 Rattus no  |
| 34         | 57    | 4.5   | 115553 | 2  | AC126077  | AC126077 Rattus no  |
| 35         | 56.8  | 4.4   | 73041  | 2  | AC127648  | AC127648 Rattus no  |
| 36         | 56.8  | 4.4   | 146431 | 2  | AC109670  | AC109670 Rattus no  |
| 37         | 56.6  | 4.4   | 157448 | 2  | AC131139  | AC131139 Rattus no  |
| 38         | 56.4  | 4.4   | 136788 | 2  | AC117044  | AC117044 Rattus no  |
| 39         | 56.2  | 4.4   | 159677 | 2  | AC113802  | AC113802 Rattus no  |
| 40         | 56.2  | 4.4   | 184402 | 2  | AC127041  | AC127041 Rattus no  |
| 41         | 55.8  | 4.4   | 179608 | 2  | AC128497  | AC128497 Rattus no  |
| 42         | 55.6  | 4.4   | 991    | 11 | PM121128  | PM121128 Homo sapi  |
| 43         | 55.6  | 4.4   | 134366 | 2  | AC119661  | AC119661 Rattus no  |
| 44         | 55.6  | 4.4   | 136014 | 2  | AC1126531 | AC1126531 Rattus no |
| 45         | 55.6  | 4.4   | 151349 | 10 | AL355176  | AL355176 Mouse DNA  |

## ALIGNMENTS

|            |   |             |     |        |                 |
|------------|---|-------------|-----|--------|-----------------|
| RESULT 1   | AL671880  | 20535 bp    | DNA | Linear | ROD 09-AUG-2002 |
| LOCUS      | AL671880  |             |     |        |                 |
| DEFINITION | Mouse DNA sequence from clone RP23-384C22 on chromosome X, complete sequence.   |             |     |        |                 |
| ACCESSION  | AL671880  |             |     |        |                 |
| VERSION    | AL671880.14   | GI:22204349 |     |        |                 |
| KEYWORDS   | HTG.  |             |     |        |                 |
| SOURCE     | house mouse.  |             |     |        |                 |
| ORGANISM   | Mus musculus  |             |     |        |                 |
| REFERENCE  | Ekakarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |             |     |        |                 |
| AUTHORS    | 1 (bases 1 to 20535)  |             |     |        |                 |
| TITLE      | Van Helmond, Z.   |             |     |        |                 |
|            | Direct Submission   |             |     |        |                 |









```

* 11879 11978: gap of 100 bp
* 11979 12686: contig of 708 bp in length
* 12687 12786: gap of 100 bp
* 12787 13495: contig of 709 bp in length
* 13496 13595: gap of 100 bp
* 13596 14306: contig of 711 bp in length
* 14307 14406: gap of 100 bp
* 14407 15103: contig of 697 bp in length
* 15104 15203: gap of 100 bp
* 15204 15921: contig of 718 bp in length
* 15922 16021: gap of 100 bp
* 16022 16696: contig of 675 bp in length
* 16697 16796: gap of 100 bp
* 16797 17519: contig of 723 bp in length
* 17520 17619: gap of 100 bp
* 17620 18337: contig of 718 bp in length
* 18338 18437: gap of 100 bp
* 18438 19132: contig of 695 bp in length
* 19133 19232: gap of 100 bp
* 19233 19906: contig of 674 bp in length
* 19907 20006: gap of 100 bp
* 20007 20705: contig of 699 bp in length
* 20706 20805: gap of 100 bp
* 20806 21498: contig of 693 bp in length
* 21499 21598: gap of 100 bp
* 21599 22302: contig of 704 bp in length
* 22303 22402: gap of 100 bp
* 22403 23108: contig of 706 bp in length
* 23109 23208: gap of 100 bp
* 23209 23927: contig of 719 bp in length
* 23928 24027: gap of 100 bp
* 24028 24748: contig of 721 bp in length
* 24749 24848: gap of 100 bp
* 24849 25566: contig of 718 bp in length
* 25567 25666: gap of 100 bp
* 25667 26370: contig of 704 bp in length
* 26371 26470: gap of 100 bp
* 26471 27168: contig of 699 bp in length
* 27170 27269: gap of 100 bp
* 27270 27963: contig of 694 bp in length
* 27964 28063: gap of 100 bp
* 28064 28757: contig of 694 bp in length
* 28758 28857: gap of 100 bp
* 28858 29570: contig of 713 bp in length
* 29571 29670: gap of 100 bp
* 29671 30383: contig of 713 bp in length
* 30384 30483: gap of 100 bp
* 30484 31189: contig of 706 bp in length
* 31190 31289: gap of 100 bp
* 31289 32018: contig of 729 bp in length
* 32019 32118: gap of 100 bp
* 32119 32843: contig of 725 bp in length
* 32844 32943: gap of 100 bp
* 32944 33614: contig of 671 bp in length
* 33615 33714: gap of 100 bp
* 33715 34436: contig of 722 bp in length
* 34437 34536: gap of 100 bp
* 34537 35257: contig of 721 bp in length
* 35258 35357: gap of 100 bp
* 35358 36063: contig of 706 bp in length
* 36064 36163: gap of 100 bp
* 36164 36856: contig of 693 bp in length
* 36857 36956: gap of 100 bp
* 36957 37661: contig of 705 bp in length
* 37662 37761: gap of 100 bp
* 37762 38464: contig of 703 bp in length
* 38465 38564: gap of 100 bp
* 38565 39269: contig of 705 bp in length
* 39270 39369: gap of 100 bp
* 39370 40052: contig of 683 bp in length
* 40053 40152: gap of 100 bp
* 40153 40863: contig of 711 bp in length
* 40864 40963: gap of 100 bp

```

```

* 40964 41682: contig of 719 bp in length
* 41683 41782: gap of 100 bp
* 41783 42490: contig of 708 bp in length
* 42491 42590: gap of 100 bp
* 42591 43311: contig of 721 bp in length
* 43312 43411: gap of 100 bp
* 43412 44137: contig of 726 bp in length
* 44138 44237: gap of 100 bp
* 44238 44946: contig of 709 bp in length
* 44947 45046: gap of 100 bp
* 45047 45758: contig of 712 bp in length
* 45759 45858: gap of 100 bp
* 45859 46553: contig of 695 bp in length
* 46554 46653: gap of 100 bp
* 46654 47366: contig of 713 bp in length
* 47367 47466: gap of 100 bp
* 47467 48179: contig of 713 bp in length
* 48180 48279: gap of 100 bp
* 48280 48944: contig of 665 bp in length
* 48945 49044: gap of 100 bp
* 49045 49763: contig of 719 bp in length
* 49764 49863: gap of 100 bp
* 49864 50570: contig of 707 bp in length
* 50571 50670: gap of 100 bp
* 50671 51392: contig of 722 bp in length
* 51393 51492: gap of 100 bp
* 51493 52184: contig of 692 bp in length
* 52185 52284: gap of 100 bp
* 52285 52992: contig of 708 bp in length
* 52993 53092: gap of 100 bp
* 53093 53787: contig of 695 bp in length
* 53788 53887: gap of 100 bp
* 53888 54595: contig of 708 bp in length
* 54596 54695: gap of 100 bp
* 54696 55404: contig of 709 bp in length
* 55405 55504: gap of 100 bp

```

Query Match 51.5%: Score 657.4; DB 2; Length 67535;  
 Best Local Similarity 86.8%; Pred.No.2,8e-125;  
 Matches 669; Conservative 0; Mismatches 101; Indels 1; Gaps 1;

```

QY 29 TGTTCGGCGGAGCAGCTTCACCCCTGCGCGCATCTTATCCGACGACCAAGCCG 88
DB 18464 TGTTCGGCGGAGCAGCTTCACCCCTGCGCGCATCTTATCCGACGACCAAGCCG 18533
QY 89 GGGATGGCTGAGGTTGGAGCCAGGAGGCTCCGACCCCGCCCTCTTCCTCCCG 148
DB 18524 GGGATGGCTGAGGTTGGAGCCAGGAGGCTCCGACCCCGCCCTCTTCCTCCCG 18583
QY 149 GCTACAGACAGCTCATGCGCGGAGTAGCTGACAGCCACGACCGGACAGCTCATGG 208
DB 18584 GCTACAGACAGCTCATGCGCGGAGTAGCTGACAGCCACGACCGGACAGCTCATGG 18643
QY 209 CCTGCTGTGCGGAGTGTCCCGGTGCGTCAAGCAGCCGTGCGTCAAGTGA 268
DB 18644 CCTGCTGTGCGGAGTGTCCCGGTGCGTCAAGCAGCCGTGCGTCAAGTGA 18703
QY 269 ACCGCGCGCGGAGCGCTGCGTCAAGTGTCAAGCAGCCGTGCGTCAAGTGA 328
DB 18704 ACCGCGCGCGGAGCGCTGCGTCAAGTGTCAAGCAGCCGTGCGTCAAGTGA 18763
QY 329 GGTGCGGAGCGCGGAGCGCGGATGCGTCAAGCAGCCGTGCGTCAAGTGA 388
DB 18764 GGTGCGGAGCGCGGAGCGCGGATGCGTCAAGCAGCCGTGCGTCAAGTGA 18823
QY 389 CCGGAGATCCCGCGCATCTGCGACAGCTGACCCCGTTCGTCGTCGACCTTCTGTG 448
DB 18824 CCGGAGATCCCGCGCATCTGCGACAGCTGACCCCGTTCGTCGTCGACCTTCTGTG 18883
QY 449 GCCTCTCTCTCTACATGAGGTTGCGGAGGAGGAGGACGACCCAGAGAGAGGAGGA 508
DB 18884 GCCTCTCTCTCTACATGAGGTTGCGGAGGAGGAGGACGACCCAGAGAGAGGAGGA 18943

```

OY 509 GCCCGCATCTCTGGGGAGCCCGGAAACCGAGAGAGTGCCCGCAGAGAAACCG 568  
 DB 18944 GCCCGCATCTCTGGGGAGCCCGGAAACCGAGAGAGTGCCCGCAGAGAAACCG 19003  
 OY 569 TCCCGCAGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627  
 DB 19004 TCCCGCAGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19063  
 OY 628 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 687  
 DB 19064 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 19123  
 OY 688 GAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 747  
 DB 19124 GAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 19183  
 OY 748 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 798  
 DB 19184 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 19234  
 RESULT 4  
 AC125993  
 LOCUS  
 DEFINITION Rattus norvegicus clone CH230-74L11, \*\*\* SEQUENCING IN PROGRESS  
 AC125993  
 VERSION AC125993.1 GI:21671835  
 KEYWORDS HTG: HTGS\_PHASE1.  
 SOURCE Rattus norvegicus.  
 ORGANISM Rattus norvegicus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 173007)  
 Muzny,D.M., Adams,C., Adlo-Ogunola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralungue,H.C., Are,J.R., Ayale,M., Banks,T., Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carmon,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Dayila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Huijck,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loussegod,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Mettler,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwaku,S., Oguh,M., Okunodu,G., Otagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Qulies,M., Ren,Y., Rivers,M., Rojas,A., Rojubenkan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,R., Thomas,S., Umanli,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,

COMMENT  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

Mu.C., Wu.Y., Wu.Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 173007)  
 Worley,K.C.  
 Direct Submission  
 Submitted (02-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 173007)  
 Worley,K.C.  
 Direct Submission  
 Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: G61Z  
 Center clone name: CH230-74L11  
 ----- Summary Statistics  
 Sequencing vector: plasmid:  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 119750 bases at least Q40  
 Consensus quality: 129093 bases at least Q30  
 Consensus quality: 135173 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drafter\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drafter_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 59 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1  
 1527: contig of 1527 bp in length  
 1627: gap of unknown length  
 1628: contig of 1153 bp in length  
 2781: gap of unknown length  
 2880: gap of unknown length  
 2881: contig of 1028 bp in length  
 3909: gap of unknown length  
 4009: contig of 1104 bp in length  
 5112: gap of unknown length  
 5212: gap of unknown length  
 5213: contig of 1464 bp in length  
 6676: gap of unknown length  
 6777: contig of 1661 bp in length  
 8437: gap of unknown length  
 8438: contig of 1042 bp in length  
 9579: gap of unknown length  
 9580: contig of unknown length  
 9680: contig of 1872 bp in length  
 11552: gap of unknown length  
 11552: contig of 1444 bp in length  
 11652: gap of unknown length  
 13095: gap of unknown length  
 13195: contig of 1336 bp in length  
 13196: gap of unknown length  
 14531: contig of 1141 bp in length  
 14532: gap of unknown length  
 15772: contig of 1829 bp in length  
 15773: gap of unknown length  
 15873: contig of 1829 bp in length  
 17701: gap of unknown length  
 17702: contig of 1101 bp in length  
 17802: gap of unknown length  
 18902: contig of 2096 bp in length  
 18903: gap of unknown length  
 19003: contig of 2096 bp in length  
 21098: gap of unknown length  
 22579: contig of 1381 bp in length  
 22579: gap of unknown length  
 22580: contig of 1305 bp in length  
 22680: contig of 1305 bp in length

```

* 23985 24084: gap of unknown length
* 24085 25116: contig of 1032 bp in length
* 25117 25216: gap of unknown length
* 25217 27027: contig of 1811 bp in length
* 27028 27128: gap of unknown length
* 27129 29170: contig of 2043 bp in length
* 29171 29270: gap of unknown length
* 31496 31595: contig of 2225 bp in length
* 31596 34124: contig of 2529 bp in length
* 34125 34224: gap of unknown length
* 34225 37124: contig of 2900 bp in length
* 37125 37224: gap of unknown length
* 37225 39707: contig of 2483 bp in length
* 39708 39807: gap of unknown length
* 39808 41454: contig of 1647 bp in length
* 41455 42338: contig of 1684 bp in length
* 42339 43339: gap of unknown length
* 43339 44619: contig of 1281 bp in length
* 44620 44719: gap of unknown length
* 44720 46804: gap of unknown length
* 46805 49282: contig of 2478 bp in length
* 49283 49382: gap of unknown length
* 49383 51919: contig of 2537 bp in length
* 51920 52019: gap of unknown length
* 52020 54858: contig of 2839 bp in length
* 54859 54958: gap of unknown length
* 54959 57481: contig of 2523 bp in length
* 57482 57581: gap of unknown length
* 57582 59384: contig of 1803 bp in length
* 59385 59484: gap of unknown length
* 59485 60768: contig of 1284 bp in length
* 60769 60869: gap of unknown length
* 60869 62576: contig of 1707 bp in length
* 62576 62676: gap of unknown length
* 62676 64968: contig of 2293 bp in length
* 64969 65068: gap of unknown length
* 65069 68642: contig of 3574 bp in length
* 68643 68742: gap of unknown length
* 68743 71330: contig of 2588 bp in length
* 71331 71430: gap of unknown length
* 71431 75415: contig of 3985 bp in length
* 75416 75515: gap of unknown length
* 75516 78191: contig of 2676 bp in length
* 78192 78291: gap of unknown length
* 78292 82195: contig of 3804 bp in length
* 82196 86302: gap of unknown length
* 86303 86402: contig of 4107 bp in length
* 86403 89944: gap of unknown length
* 89945 90044: contig of 3542 bp in length
* 90045 94437: gap of unknown length
* 94438 94537: contig of 4393 bp in length
* 94538 97928: gap of unknown length
* 97929 98028: contig of 3391 bp in length
* 98029 101684: gap of unknown length
* 101685 101784: contig of 3656 bp in length
* 101785 105383: gap of unknown length
* 105384 105483: contig of 3599 bp in length
* 105484 108194: gap of unknown length
* 108195 108294: contig of 2711 bp in length
* 108295 111907: gap of unknown length
* 111908 112007: contig of 3613 bp in length
* 112009 117232: gap of unknown length
* 117233 117332: contig of 5225 bp in length
* 117333 121679: gap of unknown length
* 121680 121779: contig of 4347 bp in length
* 121780 127615: gap of unknown length
* 127616 127715: contig of 5836 bp in length
* 127716 133212: gap of unknown length
* 133213 133312: contig of 5497 bp in length

```

```

* 133313 138145: contig of 4833 bp in length
* 138146 138245: gap of unknown length
* 138246 143564: contig of 5319 bp in length
* 143565 143664: gap of unknown length

Query Match      48.28; Score 615.4; DB 2: Length 173007;
Best Local Similarity 87.88; Pred. No. 1.3e-116;
Matches 707; Conservative 0; Mismatches 91; Indels 7; Gaps 3;

QY 3 GCGCGGCGAGGCGGAGGAGCACCATTGTCGGCGGAGCAGTTCACCCCTGCCCGCA 62
Db 104075 GCGCGGCGAGGCGGAGGAGCACCATTGTCGGCGGAGCAGCACCCTGCCCGCA 104113
QY 63 TCCCTTATCCG---AGCCACCAAAAGCCGGGATGCTGAGGTTGGACGAGGGCTTG 119
Db 104135 TCCCTTATCCG---AGCCACCAAAAGCCGGGATGCTGAGGTTGGACGAGGGCTTG 104194
QY 120 CCGACCGCGCGCGCGCGCTTCCTCCGCCGCTACAGACAGCTCATGGCCGGAGTACGT 179
Db 104195 CAGCGCGGAGCGCGCGCGCTTCCTCCGCCGCTACAGACAGCTCATGGCCGGAGTACGT 104254
QY 180 CGACAGCCACACGCGGCGAGCAGCTCATGGCCGCTTCGCCGATGGTCCCGCTCGGT 239
Db 104255 TGACAGCTATCAGAGCAGCAGCAGCTCATGGCCGCTTCGCCGATGGTCCCGCTCGGT 104314
QY 240 CAGCAGCGCTGAGCGCTCGGT---GCAGGTGAACCGCGCGCGCGCGCGCTCGGTACGTT 298
Db 104315 CAGCAGCGCTGAGCGCTCGGT---GCAGGTGAACCGCGCGCGCGCGCGCTCGGTACGTT 104374
QY 299 CACTCGGCGCGCGCGCGCTCGGT---GCAGGTGAACCGCGCGCGCGCGCGCTCGGTACGTT 358
Db 104375 CGCTCGGCGCGCGCGCGCTCGGT---GCAGGTGAACCGCGCGCGCGCGCGCTCGGTACGTT 104434
QY 359 GTTCTCTCAACCCCGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 418
Db 104435 GTTCTCTCAACCCCGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104494
QY 419 TACCGCGCTTCGCGCGCTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 478
Db 104495 TACCGCGCTTCGCGCGCTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 104554
QY 479 GCAAGCGACACCCACGAGAGGAGAGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 538
Db 104555 ACAGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104614
QY 539 AGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 598
Db 104615 AGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 104674
QY 599 TTCAGCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658
Db 104675 TCGAGCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 104731
QY 659 TCGCGCGCGATGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 718
Db 104732 TCGCGCGCGATGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 104791
QY 719 AGGACCGCGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 778
Db 104792 AGGACCGCGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 104851
QY 779 ACAAGGAGGCGCTGCTTTCAGATT 803
Db 104852 ACAAGGAGGCGCTGCTTTCAGAGT 104876

RESULT 5
AC127083/c AC127083 89551 bp DNA linear HTG 31-JUL-2002
LOCUS Rattus norvegicus clone CH230-69P8, *** SEQUENCING IN PROGRESS ***
DEFINITION 48 unordered pieces.
ACCESSION AC127083
VERSION AC127083.1 GI:21743720

```

KEYWORDS  
SOURCE  
ORGANISM

HTG: HTGS, PHASE1.  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 89551)  
Munzy, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Albrooks, S.L., Amaralunga, H.C., Are, J.R., Ayale, M., Banks, T.,  
Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowie, S., Brivea, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Dayla, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delgado, K.R., Delgado, O., Denn, A.L., Ding, Y., Dihn, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Einhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, F., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, R.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,  
Homs, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtenberg, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mathew, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Mosier, N., Neal, D., Newton, J., Newton, N., Nguyen, N., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G.,  
Oranuy, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peterson, L., Pickens, R., Pilms, E., Pu, L.L., Qules, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherer, S., Scott, G., Shen, H., Shoshitani, N., Sisson, I.,  
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vanson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wiczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

TITLE  
JOURNALREFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 89551)  
Worley, K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 89551)  
Worley, K.C.  
Direct Submission  
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: KAMV  
Center clone name: CH230-69F8  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329

Consensus quality: 42669 bases at least Q40  
Consensus quality: 46367 bases at least Q30  
Consensus quality: 48722 bases at least Q20

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 48 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1151: contig of 1151 bp in length  
1152 1251: gap of unknown length  
1252 2322: contig of 1071 bp in length  
2323 2422: gap of unknown length  
2423 3692: contig of 1270 bp in length  
3693 3792: gap of unknown length  
3793 5136: contig of 1344 bp in length  
5137 5236: gap of unknown length  
5237 6325: contig of 1089 bp in length  
6326 6425: gap of unknown length  
6426 8049: contig of 1624 bp in length  
8050 8149: gap of unknown length  
8150 9419: contig of 1270 bp in length  
9420 9519: gap of unknown length  
9520 10641: contig of 1122 bp in length  
10642 10741: gap of unknown length  
10742 11933: contig of 1192 bp in length  
11934 12033: gap of unknown length  
12034 13462: contig of 1329 bp in length  
13463 13462: gap of unknown length  
13463 14779: contig of 1317 bp in length  
14779 14879: gap of unknown length  
14880 16451: contig of 1572 bp in length  
16452 17915: gap of unknown length  
17916 18015: contig of 1364 bp in length  
18016 19444: gap of unknown length  
19445 19544: contig of 1428 bp in length  
19545 20626: gap of unknown length  
20627 20726: contig of 1083 bp in length  
20727 22100: gap of unknown length  
22101 22200: contig of 1374 bp in length  
22201 23252: gap of unknown length  
23253 23352: contig of 1052 bp in length  
23353 23352: gap of unknown length  
23353 24945: contig of 1593 bp in length  
24946 25045: gap of unknown length  
25046 26886: contig of 1841 bp in length  
26887 28215: gap of unknown length  
28216 28315: contig of 1229 bp in length  
28316 28315: gap of unknown length  
28316 29904: contig of 1589 bp in length  
29905 30004: gap of unknown length  
30005 31803: contig of 1799 bp in length  
31804 31903: gap of unknown length  
31904 33938: contig of 2035 bp in length  
33939 34038: gap of unknown length  
34039 35078: contig of 1040 bp in length  
35079 35178: gap of unknown length  
35179 37353: contig of 2174 bp in length  
37354 37452: gap of unknown length  
37453 39726: contig of 2274 bp in length  
39727 39825: gap of unknown length  
39826 41535: contig of 1729 bp in length  
41536 41555: gap of unknown length  
41556 43271: contig of 1616 bp in length  
43272 43372: gap of unknown length  
43373 44785: contig of 1414 bp in length  
44786 44885: gap of unknown length  
44886 46567: contig of 1682 bp in length



```

* 46568 46667: gap of unknown length
* 46668 49449: contig of 2782 bp in length
* 49450 49549: gap of unknown length
* 49550 51783: contig of 2234 bp in length
* 51784 51883: gap of unknown length
* 51884 54046: contig of 2163 bp in length
* 54047 54146: gap of unknown length
* 54147 56182: contig of 2036 bp in length
* 56183 56283: gap of unknown length
* 56283 57934: contig of 1652 bp in length
* 57935 58035: gap of unknown length
* 58035 59825: contig of 1791 bp in length
* 59826 59926: gap of unknown length
* 59926 61535: contig of 1610 bp in length
* 61536 61635: gap of unknown length
* 61635 63473: contig of 1838 bp in length
* 63474 63573: gap of unknown length
* 63573 66376: contig of 2803 bp in length
* 66377 66476: gap of unknown length
* 66477 68641: contig of 2165 bp in length
* 68642 70180: gap of unknown length
* 70181 70280: contig of 1439 bp in length
* 70281 72432: gap of unknown length
* 72433 72533: gap of unknown length
* 72533 73917: contig of 1385 bp in length
* 73918 74017: gap of unknown length
* 74018 77137: contig of 3120 bp in length
* 77138 77237: gap of unknown length
* 77238 79997: contig of 2760 bp in length
* 79998 80097: gap of unknown length
* 80098 82859: contig of 2762 bp in length
* 82860 82959: gap of unknown length
* 82960 85680: contig of 2721 bp in length
* 85681 85780: gap of unknown length
* 85781 89551: contig of 3771 bp in length.

```

FEATURES  
SOURCE  
Location/Qualifiers  
1. 89551

/Organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-69P8"

BASE COUNT 23140 a 19260 c 19626 g 22768 t 4757 others

ORIGIN

Query Match 45.8% Score 585.2; DB 2: Length 89551;  
Best Local Similarity 87.6% Pred. No. 2,1e-110;  
Matches 664: Conservative 0; Mismatches 88; Indels 6; Gaps 2;

```

OY 49 CACCCCTGCCCATCTTATCCGC---AGGCCAACAAAGCCGGGATGCTGAGAGTTTC 105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39687 CACCCATGCCCATCTTATCCGCCTCTGTGACCAAGCCGGGATGCTGAGAGTTT 39628
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 106 GAGCCAGGGGCTGCGAGCCGGCCGCTCTCCCTCCCGGTACAGACACTCATG 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39627 GAGCCAGGGGCTGCGAGCCGGCCGCTCTCCCTCCCGGTACAGACACTCATG 39368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 166 GCCGGGAGTACGTGACAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39567 GCCGGGAGTACTTGTACAGCATGACAGGAGGAGGAGGAGGAGGAGGAGGAG 39508
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 226 GGTCCCGCGGTGCTGAGACAGCCGCTGCGGTGCGGTGCGGTGCGGTGCGGTG 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39507 GGTCCCGCGGTGCTGAGACAGCCGCTGCGGTGCGGTGCGGTGCGGTGCGGT 39448
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 286 TCGGTGAGTGTACTGACGTGGGCGCGACGTGACGTGACGTGACGTGACGTGAC 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39447 TCGGTGAGTGTACTGACGTGGGCGCGACGTGACGTGACGTGACGTGACGTG 39388
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 346 GAGCCGCGCATGCGGTCTCTGTCAACCCCGGTGCGGTGCGGTGCGGTGCGGTG 405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39387 GAGCCGCGCATGCGGTCTCTGTCAACCCCGGTGCGGTGCGGTGCGGTGCGGT 39328
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 406 TCTGTGACAGACCGTACGCCCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 465
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39327 TCTGTGCGCACCGTCCGCCCTTACTCCGCCCTGACCTTCCGTTGCTCTCTCCCTG 39268
OY 466 GAGTTGGGGAGGCGCAGACACCCAGCAAGAGGAGGAGGAGGAGGAGGAGGAGG 535
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39267 GAGTTGGGGAGGCGCAGACACCCAGCAAGAGGAGGAGGAGGAGGAGGAGGAGG 39208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 526 ACCGGGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39207 ACCGGGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 39148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 586 GAGGAGGCGCATGTTTACGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39147 GAGGAGGCGCATGTTTACGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 39091
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 646 GAGCGGAACAGTGTGCGCGCGCATGAGTGTGAGCTTGGAGCGAGGAGGAGGAGG 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39090 GAGCGGAACAGTGTGCGCGCGCATGAGTGTGAGCTTGGAGCGAGGAGGAGGAGG 39031
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 706 GAGGAGTGGCTGAGACCGCGGTGATTGCGATGCGCTTGGAGCGAGGAGGAGGAG 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39030 GTCGAGATGGCTGAGACCGCGGTGATTGCGATGCGCTTGGAGCGAGGAGGAGGAG 38971
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 766 AGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 803
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 38970 AGCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38933
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 6  
AC126519/c 169908 bp DNA linear HTG 24-JUL-2002  
DEFINITION  
Rattus norvegicus clone CH230-159N5, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 60 unordered pieces.

AC126519  
AC126519.1 GI:21700481  
VERSION  
HTG: HTGS\_PHASE1.  
KEYWORDS  
Rattus norvegicus.  
SOURCE  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 169908)  
Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barberia,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkelt,C., Burrell,R.L., Byrd,N.C.,  
Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M., Carroll,L., Dederich,D.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Delaney,K.R.,  
Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisl,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hoggins,M., Holloway,C., Hollins,B.,  
Homs,I.F., Howard,S., Huber,J., Hult,K., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolyvet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Liu,C., Liu,J., Liu,W., Loulsegel,H.,  
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,G., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Nohbbat,K., Morgan,M., Morris,S.,  
Nguyen,N., Nickerson,E., Nockenhoe,S., Ogih,M., Okunou,G.,  
Oragunye,N., Oyiedo,R., Pace,A., Payton,B., Peety,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Roife,M., Ruiz,S., Savery,G.,

Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swalek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Taylor, J., Taylor, C., Taylor, T., Tellro, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 169908)  
Worley, K.C.  
Direct Submission  
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 169908)  
Worley, K.C.  
Direct Submission  
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GZCZ  
Center clone name: CH230-159N5  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap, version 0.990329  
Consensus quality: 125808 bases at least Q40  
Consensus quality: 132708 bases at least Q30  
Consensus quality: 136844 bases at least Q20

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Gendbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Gendbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 60 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N. But the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1574: contig of 1574 bp in length  
1575 1674: gap of unknown length  
1675 2679: contig of 1005 bp in length  
2680 2779: gap of unknown length  
2780 3878: contig of 1099 bp in length  
3879 3978: gap of unknown length  
3979 5423: contig of 1445 bp in length  
5424 5523: gap of unknown length  
5523 6788: contig of 1265 bp in length  
6789 6888: gap of unknown length  
6889 8285: contig of 1397 bp in length  
8286 8385: gap of unknown length  
8386 9590: contig of 1205 bp in length  
9591 10999: contig of 1309 bp in length  
10999 11000: gap of unknown length  
11000 11009: gap of 1399 bp in length  
11009 12498: contig of 1399 bp in length  
12498 12598: gap of unknown length  
12599 14147: contig of 1549 bp in length  
14148 14247: gap of unknown length  
14248 16266: contig of 2019 bp in length  
16267 16366: gap of unknown length  
16367 18233: contig of 1867 bp in length  
18234 18333: gap of unknown length

18334 19761: contig of 1428 bp in length  
19762 19861: gap of unknown length  
19862 21637: contig of 1776 bp in length  
21638 21737: gap of unknown length  
21738 23243: contig of 1406 bp in length  
23244 25296: gap of unknown length  
25297 25396: contig of 2053 bp in length  
25397 25397: gap of unknown length  
25398 26534: contig of 1137 bp in length  
26535 26634: gap of unknown length  
26635 28279: contig of 1645 bp in length  
28280 28379: gap of unknown length  
28380 30343: contig of 1965 bp in length  
30344 30443: gap of unknown length  
30444 32122: contig of 1678 bp in length  
32123 32222: gap of unknown length  
32223 34582: contig of 2361 bp in length  
34583 34682: gap of unknown length  
34683 35980: contig of 1298 bp in length  
35981 36081: gap of unknown length  
36082 37395: contig of 1314 bp in length  
37396 37495: gap of unknown length  
37496 39068: contig of 1474 bp in length  
39069 39069: gap of unknown length  
39070 40199: contig of 1131 bp in length  
40200 40299: gap of unknown length  
40300 41804: contig of 1505 bp in length  
41805 41905: gap of unknown length  
41906 43570: contig of 1665 bp in length  
43571 43670: gap of unknown length  
43671 45680: contig of 2011 bp in length  
45681 45780: gap of unknown length  
45781 48216: contig of 2435 bp in length  
48217 48316: gap of unknown length  
48317 51069: contig of 2753 bp in length  
51070 51169: gap of unknown length  
51170 53667: contig of 2498 bp in length  
53668 53767: gap of unknown length  
53768 55833: contig of 2066 bp in length  
55834 55932: gap of unknown length  
55933 55933: contig of 1628 bp in length  
55934 57660: gap of unknown length  
57661 57661: gap of unknown length  
57662 59004: contig of 2244 bp in length  
59005 60005: gap of unknown length  
60006 62974: contig of 2970 bp in length  
62975 63075: gap of unknown length  
63076 66492: contig of 3417 bp in length  
66493 66591: gap of unknown length  
66592 68271: contig of 1680 bp in length  
68272 68372: gap of unknown length  
68373 70763: contig of 2391 bp in length  
70764 70863: gap of unknown length  
70864 73263: contig of 2401 bp in length  
73264 73363: gap of unknown length  
73364 75903: gap of 3540 bp in length  
75904 77003: gap of unknown length  
77004 79444: contig of 2441 bp in length  
79445 79544: gap of unknown length  
79545 81957: contig of 2413 bp in length  
81958 82058: gap of unknown length  
82059 85214: contig of 3157 bp in length  
85215 85314: gap of unknown length  
85315 87778: contig of 2464 bp in length  
87779 87879: gap of unknown length  
87880 90897: contig of 3019 bp in length  
90898 90997: gap of unknown length  
90998 96387: contig of 5390 bp in length  
96388 96487: gap of unknown length  
96488 100595: contig of 4108 bp in length  
100596 100696: gap of unknown length  
100697 104991: contig of 4296 bp in length  
104992 105092: gap of unknown length  
105093 109386: contig of 4295 bp in length

\* 109387 109486: gap of unknown length  
 \* 109487 113734: contig of 4248 bp in length  
 \* 113735 113834: gap of unknown length  
 \* 113835 116884: contig of 3050 bp in length  
 \* 116885 116984: gap of unknown length  
 \* 116985 121008: contig of 4024 bp in length  
 \* 121009 121108: gap of unknown length  
 \* 121109 124634: contig of 3526 bp in length  
 \* 124635 124735: gap of unknown length  
 \* 124735 130307: contig of 5573 bp in length  
 \* 130308 130407: gap of unknown length

Query Match 35.4% Score 452.4: DB 2: Length 169908;  
 Best Local Similarity 87.5% Pred. No. 4.9e-83;  
 Matches 530: Conservative 0: Mismatches 71: Indels 5: Gaps 3:

QY 199 CAGCTCATGCGCCCTGCTCTGCTGCGGATGGGTGCGCGGTG-GTCAAGACCCGTGACGCTGC 257  
 DB 10971 CAGCTCATGCGCCCTGCTCTGCTGCGGATGGGTGCGCGGTGCGGTCACAGCCGCGACGCTGC 10912  
 QY 258 GGTGAGGTAACCCGCGCGCGGAGCGGCTGCGGTGCGGTCACGTCGCGGCGCGCGCGCGCT 317  
 DB 10911 GGTGAGGTAACCCGCGCGCGGAGCGGCTGCGGTGCGGTCACGTCGCGGCGCGCGCGCT 10852  
 QY 318 GCAGCCTGCGAGGTCGCGAGCCGCGAGCCCGCGATCGGGTCTCTGTCACCCCGCTGCG 377  
 DB 10851 GCAGCCTGCGAGGTCGCGAGCCGCGAGCCCGCGATCGGGTCTCTGTCACCCCGCTGCG 10792  
 QY 378 CCAGCCCGCGCGCGGAGATCCCGCGATCTCTGCGAGACCGTACCGCCGCTGCTGCTCCGT 437  
 DB 10791 CCAGCCCGCGCGCGGAGATCCCGCGATCTCTGCGAGACCGTACCGCCGCTGCTGCTCCGT 10732  
 QY 438 GACCTTCTGTGCT 497  
 DB 10731 GACCTTCTGTGCT 10672  
 QY 498 GGGAGAGGGAGACCCCGCGATCTCTGCGGAGACCCCGAGCCGAGAGAGTGTGCGCGC 557  
 DB 10671 -GGAAGAGGGAGACCCCGCGATCTCTGCGGAGACCCCGAGCCGAGAGAGTGTGCGCGC 10613  
 QY 558 GAGGAACCGGTCCT 617  
 DB 10612 GATGAACCGGTCCT 10553  
 QY 618 CGGCTGGAGAGACGACCAACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 677  
 DB 10552 TG---GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10496  
 QY 678 GCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 737  
 DB 10495 GCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10436  
 QY 738 TGCCCTCGAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 797  
 DB 10435 TGCCCTCGAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10376  
 QY 798 CCAAGT 803  
 DB 10375 CCAGGT 10370

RESULT 7  
 AC108848  
 LOCUS AC108848 212848 bp DNA linear HTG 28-APR-2002  
 DEFINITION Mus musculus clone RP23-300K5, WORKING DRAFT SEQUENCE, 36 ordered  
 pieces.  
 ACCESSION AC108848  
 VERSION AC108848.2 GI:20336129  
 KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Scleroglossa; Muridae; Murinae; Mus.

# REFERENCE

1 (bases 1 to 212848)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Mus musculus, clone RP23-300K5  
 Unpublished  
 2 (bases 1 to 212848)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,  
 Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,  
 Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,  
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,  
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gargana, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Lachowicz, K., Lamazares, R.,  
 Landers, T., Lenocksky, J., Levine, R., Liu, G., Maclean, C.,  
 Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,  
 McEwan, P., McKernan, K., Melrim, J., Menus, L., Mihova, T.,  
 Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,  
 Reta, R., Ribbeck, M., Riley, R., Rise, C., Rogov, P., Roman, J.,  
 Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,  
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Watson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

# TITLE

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 212848)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
 Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,  
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gargana, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Lachowicz, K., Lamazares, R.,  
 Landers, T., Lenocksky, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., Melrim, J., Menus, L.,  
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Reta, R., Ribbeck, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Watson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

# TITLE

Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 28, 2002 this sequence version replaced gi:18450109.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

# COMMENT

Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu

# Project Information

Center project name: L21014  
 Center clone name: 300\_K\_5  
 Summary Statistics

# Sequencing vector

Plasmid: n/a; 100% of reads  
 Chemistry: dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 200069 bases at least Q40

Consensus quality: 206302 bases at least Q30  
 Consensus quality: 208218 bases at least Q20  
 Insert size: 209348; sum-of-contigs  
 Quality coverage: 5.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

1 641: contig of 641 bp in length  
 642 741: gap of 100 bp  
 742 1427: contig of 686 bp in length  
 1428 1527: gap of 100 bp  
 1528 3170: contig of 1643 bp in length  
 3171 3270: gap of 100 bp  
 3271 4396: contig of 1126 bp in length  
 4397 4496: gap of 100 bp  
 4497 5909: contig of 1433 bp in length  
 5910 6009: gap of 100 bp  
 6010 7583: contig of 1574 bp in length  
 7584 7683: gap of 100 bp  
 7684 8975: contig of 1252 bp in length  
 8976 9075: gap of 100 bp  
 9076 10257: contig of 1182 bp in length  
 10258 10357: gap of 100 bp  
 10359 11608: contig of 1251 bp in length  
 11609 11708: gap of 100 bp  
 11709 13289: contig of 1581 bp in length  
 13290 13389: gap of 100 bp  
 13390 15335: contig of 1946 bp in length  
 15336 15435: gap of 100 bp  
 15436 16250: contig of 815 bp in length  
 16251 16350: gap of 100 bp  
 16351 17669: contig of 1319 bp in length  
 17670 17769: gap of 100 bp  
 17770 19907: contig of 2138 bp in length  
 19908 20007: gap of 100 bp  
 20008 21470: contig of 1463 bp in length  
 21471 21570: gap of 100 bp  
 21571 23086: contig of 1516 bp in length  
 23087 23186: gap of 100 bp  
 23187 24723: contig of 1537 bp in length  
 24724 24823: gap of 100 bp  
 24824 27746: contig of 2923 bp in length  
 27747 27846: gap of 100 bp  
 27847 29873: contig of 2027 bp in length  
 29874 29973: gap of 100 bp  
 29974 33212: contig of 3239 bp in length  
 33213 33312: gap of 100 bp  
 33313 61732: contig of 28420 bp in length  
 61733 61832: gap of 100 bp  
 61833 63792: contig of 1960 bp in length  
 63793 63892: gap of 100 bp  
 63893 67231: contig of 3339 bp in length  
 67232 67331: gap of 100 bp  
 67332 70922: contig of 3591 bp in length  
 70923 71022: gap of 100 bp  
 71023 76930: contig of 5908 bp in length  
 76931 77030: gap of 100 bp  
 77031 82623: contig of 5593 bp in length  
 82624 82723: gap of 100 bp  
 82724 87875: contig of 5152 bp in length  
 87876 87975: gap of 100 bp  
 87976 93774: contig of 5799 bp in length  
 93775 93874: gap of 100 bp  
 93875 104205: contig of 10331 bp in length  
 104206 104305: gap of 100 bp  
 104306 119104: contig of 14799 bp in length

## FEATURES

119105 119204: gap of 100 bp  
 119205 132299: contig of 13095 bp in length  
 132300 132399: gap of 100 bp  
 132400 148970: contig of 16571 bp in length  
 148971 149070: gap of 100 bp  
 149071 165714: contig of 16644 bp in length  
 165715 165814: gap of 100 bp  
 165815 183369: contig of 17555 bp in length  
 183370 183469: gap of 100 bp  
 183470 210291: contig of 26822 bp in length  
 210292 210391: gap of 100 bp  
 210392 212848: contig of 2457 bp in length.

## SOURCE

1. 212848  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP23-30K5"  
 /clone\_id="RP23-23 Female Mouse BAC"  
 1. 641

misc\_feature  
 742..1427  
 /note="assembly\_fragment"

misc\_feature  
 1528..3170  
 /note="assembly\_fragment"

misc\_feature  
 3271..4396  
 /note="assembly\_fragment"

misc\_feature  
 4497..5909  
 /note="assembly\_fragment"

misc\_feature  
 6010..7583  
 /note="assembly\_fragment"

misc\_feature  
 7684..8975  
 /note="assembly\_fragment"

misc\_feature  
 9076..10257  
 /note="assembly\_fragment"

misc\_feature  
 10358..11608  
 /note="assembly\_fragment"

misc\_feature  
 11709..13289  
 /note="assembly\_fragment"

misc\_feature  
 13390..15335  
 /note="assembly\_fragment"

misc\_feature  
 15436..16250  
 /note="assembly\_fragment"

misc\_feature  
 16351..17669  
 /note="assembly\_fragment"

misc\_feature  
 17670..19907  
 /note="assembly\_fragment"

misc\_feature  
 19908..20007  
 /note="assembly\_fragment"

misc\_feature  
 20008..21470  
 /note="assembly\_fragment"

misc\_feature  
 21471..21570  
 /note="assembly\_fragment"

misc\_feature  
 21571..23086  
 /note="assembly\_fragment"

misc\_feature  
 23087..23186  
 /note="assembly\_fragment"

misc\_feature  
 23187..24723  
 /note="assembly\_fragment"

misc\_feature  
 24724..24823  
 /note="assembly\_fragment"

misc\_feature  
 24824..27746  
 /note="assembly\_fragment"

misc\_feature  
 27747..27846  
 /note="assembly\_fragment"

misc\_feature  
 27847..29873  
 /note="assembly\_fragment"

misc\_feature  
 29874..29973  
 /note="assembly\_fragment"

misc\_feature  
 29974..33212  
 /note="assembly\_fragment"

misc\_feature  
 33213..61732  
 /note="assembly\_fragment"

misc\_feature  
 61733..61832  
 /note="assembly\_fragment"

misc\_feature  
 61833..63792  
 /note="assembly\_fragment"

## FEATURES

119105 119204: gap of 100 bp  
 119205 132299: contig of 13095 bp in length  
 132300 132399: gap of 100 bp  
 132400 148970: contig of 16571 bp in length  
 148971 149070: gap of 100 bp  
 149071 165714: contig of 16644 bp in length  
 165715 165814: gap of 100 bp  
 165815 183369: contig of 17555 bp in length  
 183370 183469: gap of 100 bp  
 183470 210291: contig of 26822 bp in length  
 210292 210391: gap of 100 bp  
 210392 212848: contig of 2457 bp in length.

## SOURCE

1. 212848  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP23-30K5"  
 /clone\_id="RP23-23 Female Mouse BAC"  
 1. 641

misc\_feature  
 742..1427  
 /note="assembly\_fragment"

misc\_feature  
 1528..3170  
 /note="assembly\_fragment"

misc\_feature  
 3271..4396  
 /note="assembly\_fragment"

misc\_feature  
 4497..5909  
 /note="assembly\_fragment"

misc\_feature  
 6010..7583  
 /note="assembly\_fragment"

misc\_feature  
 7684..8975  
 /note="assembly\_fragment"

misc\_feature  
 9076..10257  
 /note="assembly\_fragment"

misc\_feature  
 10358..11608  
 /note="assembly\_fragment"

misc\_feature  
 11709..13289  
 /note="assembly\_fragment"

misc\_feature  
 13390..15335  
 /note="assembly\_fragment"

misc\_feature  
 15436..16250  
 /note="assembly\_fragment"

misc\_feature  
 16351..17669  
 /note="assembly\_fragment"

misc\_feature  
 17670..19907  
 /note="assembly\_fragment"

misc\_feature  
 19908..20007  
 /note="assembly\_fragment"

misc\_feature  
 20008..21470  
 /note="assembly\_fragment"

misc\_feature  
 21471..21570  
 /note="assembly\_fragment"

misc\_feature  
 21571..23086  
 /note="assembly\_fragment"

misc\_feature  
 23087..23186  
 /note="assembly\_fragment"

misc\_feature  
 23187..24723  
 /note="assembly\_fragment"

misc\_feature  
 24724..24823  
 /note="assembly\_fragment"

misc\_feature  
 24824..27746  
 /note="assembly\_fragment"

misc\_feature  
 27747..27846  
 /note="assembly\_fragment"

misc\_feature  
 27847..29873  
 /note="assembly\_fragment"

misc\_feature  
 29874..29973  
 /note="assembly\_fragment"

misc\_feature  
 29974..33212  
 /note="assembly\_fragment"

misc\_feature  
 33213..61732  
 /note="assembly\_fragment"

misc\_feature  
 61733..61832  
 /note="assembly\_fragment"

misc\_feature  
 61833..63792  
 /note="assembly\_fragment"

## FEATURES

119105 119204: gap of 100 bp  
 119205 132299: contig of 13095 bp in length  
 132300 132399: gap of 100 bp  
 132400 148970: contig of 16571 bp in length  
 148971 149070: gap of 100 bp  
 149071 165714: contig of 16644 bp in length  
 165715 165814: gap of 100 bp  
 165815 183369: contig of 17555 bp in length  
 183370 183469: gap of 100 bp  
 183470 210291: contig of 26822 bp in length  
 210292 210391: gap of 100 bp  
 210392 212848: contig of 2457 bp in length.

## SOURCE

1. 212848  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP23-30K5"  
 /clone\_id="RP23-23 Female Mouse BAC"  
 1. 641

misc\_feature  
 742..1427  
 /note="assembly\_fragment"

misc\_feature  
 1528..3170  
 /note="assembly\_fragment"

misc\_feature  
 3271..4396  
 /note="assembly\_fragment"

misc\_feature  
 4497..5909  
 /note="assembly\_fragment"

misc\_feature  
 6010..7583  
 /note="assembly\_fragment"

misc\_feature  
 7684..8975  
 /note="assembly\_fragment"

misc\_feature  
 9076..10257  
 /note="assembly\_fragment"

misc\_feature  
 10358..11608  
 /note="assembly\_fragment"

misc\_feature  
 11709..13289  
 /note="assembly\_fragment"

misc\_feature  
 13390..15335  
 /note="assembly\_fragment"

misc\_feature  
 15436..16250  
 /note="assembly\_fragment"

misc\_feature  
 16351..17669  
 /note="assembly\_fragment"

misc\_feature  
 17670..19907  
 /note="assembly\_fragment"

misc\_feature  
 19908..20007  
 /note="assembly\_fragment"

misc\_feature  
 20008..21470  
 /note="assembly\_fragment"

misc\_feature  
 21471..21570  
 /note="assembly\_fragment"

misc\_feature  
 21571..23086  
 /note="assembly\_fragment"

misc\_feature  
 23087..23186  
 /note="assembly\_fragment"

misc\_feature  
 23187..24723  
 /note="assembly\_fragment"

misc\_feature  
 24724..24823  
 /note="assembly\_fragment"

misc\_feature  
 24824..27746  
 /note="assembly\_fragment"

misc\_feature  
 27747..27846  
 /note="assembly\_fragment"

misc\_feature  
 27847..29873  
 /note="assembly\_fragment"

misc\_feature  
 29874..29973  
 /note="assembly\_fragment"

misc\_feature  
 29974..33212  
 /note="assembly\_fragment"

misc\_feature  
 33213..61732  
 /note="assembly\_fragment"

misc\_feature  
 61733..61832  
 /note="assembly\_fragment"

misc\_feature  
 61833..63792  
 /note="assembly\_fragment"



| QY         | 63   | TCCTTATCCGAGGCGACAAAGCCGGGGATGGCTGAGAGTTTCGAGGCGACAGGGGCTGCCG | 122                        |
|------------|--|---|----------------------------|
| Db         | 88210  | TCCTTATCCGAGGCGACAAAGCCGGGGATGGCTGAGAGTTTCGAGGCGACAGGGGCTGCCG | 88151                      |
| QY         | 123  | ACCCGCGCCCCCTCTCTTCTCCCGGCTACAGACAGCTACTGGCCCGGGAGTACGTGCA    | 182                        |
| Db         | 88150  | ACCCGCGCCCCCTCTCTTCTCCCGGCTACAGACAGCTACTGGCCCGGGAGTACGTGCA    | 88091                      |
| QY         | 183  | CAGCCACCAAGGGGACAGCTCATGGCCCTGTCGGGAGATGGGTCCCGGTCGGTCAG      | 242                        |
| Db         | 88090  | CCGGCACCAGGGGGACAGCTCATGGCCCTGTCGGGAGATGGGTCCCGGTCGGTCAG      | 88031                      |
| QY         | 243  | CAGCCGTGAGCGCTGCGGTGCAAGTGAACCCGGCCGCGAGCGCTCGGTGCACTGTCACT   | 302                        |
| Db         | 88030  | CAGCCGTGAGCGCTGCGGTGCAAGTGAACCCGGCCGCGAGCGCTCGGTGCACTGTNNNNN  | 87971                      |
| QY         | 303  | CGGGCGCCGCAACGCTGCAGGCTGCAGGGTGGCGAGCCGCCCGACGCCCATCGGGTTTC   | 362                        |
| Db         | 87970  | NN  | 87911                      |
| QY         | 363  | CTGTCAACCCCGCTGGCCACGCCGGCGCGGAGATGCCCGGATCTGGCAGACCGTAGC     | 422                        |
| Db         | 87910  | NN  | 87851                      |
| QY         | 423  | CCCCCTTC 429  |                            |
| Db         | 87850  | ACAGTTTC 87844  |                            |
| RESULT 9   | AC127083   | 89551 bp  | DNA linear HTG 31-JUL-2002 |
| LOCUS      | AC127083   |   |                            |
| DEFINITION | Rattus norvegicus clone CH230-69F8, *** SEQUENCING IN PROGRESS *** |   |                            |

ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
1 (bases 1 to 89551)  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
Barbada, J., Benton, J., Bimaga, K., Blankenhorn, K., Bonnin, D.,  
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Deedrich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.,  
Dothanaithe, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Eamhart, C., Edgar, D., Edwards, C.C., Elhai, C., Escoto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P.,  
Gabis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorelli, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havila, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
Honsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudash, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvach, J., Kovar, C.,  
Kralovic, J., Kureshi, A., Landry, N., Deal, B., Lewis, L.C., Lewis, L.,

Lozano, R.U., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,L., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Morris,K., Nishikawa,S.

Moser, M., Neal, D., Newton, J., Newton, S. N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokweto, S., Oguh, M., Okwunnu, G., Oragunye, N., Oyiedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Peters, L., Pickens, R., Primus, E., Pu, L.L., Qulles, M., Ren, Y.,  
Rivas, M., Rojas, A., Rojibokan, I., Rolfe, S., Ruiz, S., Severy, G.,  
Scherer, S., Scott, G., Shen, H., Shoochari, N., Sisson, I.,  
Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Swalek, A., Tabor, P., Tamerisa, A., Tamerisa, R., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Uman, K., Vasquez, L., Vera, Y., Villalon, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,  
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Wetstock, G., and Gibbs, R.  
Unpublished  
2 (bases 1 to 89551)  
Morley, K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 89551)  
Morley, K.C.  
Direct Submission  
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
-----  
Center: Baylor College of Medicine  
Genome Center  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
-----  
Project Information  
Center project name: KAWY  
Center clone name: CH230-69F8  
-----  
Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 42669 bases at least Q40  
Consensus quality: 46367 bases at least Q30  
Consensus quality: 48722 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 48 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1151: contig of 1151 bp in length  
\* 1152 1251: gap of unknown length  
\* 1252 2322: contig of 1071 bp in length  
\* 2323 2422: gap of unknown length  
\* 2423 3692: contig of 1270 bp in length  
\* 3693 3792: gap of unknown length  
\* 3793 5136: contig of 1344 bp in length  
\* 5137 5236: gap of unknown length  
\* 5237 6325: contig of 1089 bp in length  
\* 6326 6425: gap of unknown length  
\* 6426 8049: contig of 1624 bp in length  
\* 8050 8149: gap of unknown length  
\* 8150 9419: contig of 1270 bp in length  
\* 9420 9519: gap of unknown length  
\* 9520 10641: contig of 1122 bp in length  
\* 10642 10741: gap of unknown length  
\* 10742 11933: contig of 1192 bp in length  
\* 11934 12033: gap of unknown length  
\* 12034 13362: contig of 1329 bp in length  
\* 13363 13462: gap of unknown length  
\* 13463 14779: contig of 1317 bp in length  
\* 14780 14879: gap of unknown length  
14880 16451: contig of 1572 bp in length  
\* 16452 16551: gap of unknown length  
\* 16552 17915: contig of 1364 bp in length  
\* 17916 18015: gap of unknown length  
\* 18016 19443: contig of 1428 bp in length  
\* 19444 19543: gap of unknown length  
\* 19544 20627: contig of 1083 bp in length  
\* 20628 20726: gap of unknown length  
\* 20727 22100: contig of 1374 bp in length  
\* 22101 22200: gap of unknown length  
\* 22201 23352: contig of 1052 bp in length  
\* 23353 23352: gap of unknown length  
\* 23353 24945: contig of 1593 bp in length  
\* 24946 25045: gap of unknown length  
\* 25046 26886: contig of 1841 bp in length  
\* 26887 26986: gap of unknown length  
\* 26987 28215: contig of 1229 bp in length  
\* 28216 28315: gap of unknown length  
\* 28316 29904: contig of 1589 bp in length  
\* 29905 30004: gap of unknown length  
\* 30005 31803: contig of 1799 bp in length  
\* 31804 31903: gap of unknown length  
\* 31904 33938: contig of 2035 bp in length  
\* 33939 34038: gap of unknown length  
\* 34039 35078: contig of 1040 bp in length  
\* 35079 35178: gap of unknown length  
\* 35179 37352: contig of 2174 bp in length  
\* 37353 37452: gap of unknown length  
\* 37453 39726: contig of 2274 bp in length  
\* 39727 39826: gap of unknown length  
\* 39827 41555: contig of 1129 bp in length  
\* 41556 41655: gap of unknown length  
\* 41656 43271: contig of 1616 bp in length  
\* 43272 43371: gap of unknown length  
\* 43372 44785: contig of 1414 bp in length  
\* 44786 44885: gap of unknown length  
\* 44886 46567: contig of 1682 bp in length  
\* 46568 46667: gap of unknown length  
\* 46667 49449: contig of 2782 bp in length  
\* 49450 49549: gap of unknown length  
\* 49550 51783: contig of 2234 bp in length  
\* 51784 51883: gap of unknown length  
\* 51883 54046: contig of 2163 bp in length  
\* 54046 54146: gap of unknown length  
\* 54147 56182: contig of 2036 bp in length  
\* 56183 56282: gap of unknown length  
\* 56283 57934: contig of 1652 bp in length  
\* 57935 58034: gap of unknown length  
\* 58035 59825: contig of 1791 bp in length  
\* 59826 59925: gap of unknown length  
\* 59926 61535: contig of 1610 bp in length  
\* 61536 61635: gap of unknown length  
\* 61636 63473: contig of 1838 bp in length  
\* 63474 63573: gap of unknown length  
\* 63574 66376: contig of 2803 bp in length  
\* 66377 66476: gap of unknown length  
\* 66477 68641: contig of 2165 bp in length  
\* 68642 68741: gap of unknown length  
\* 68742 70180: contig of 1439 bp in length  
\* 70181 72432: gap of unknown length  
\* 72433 72532: contig of 2152 bp in length  
\* 72533 73917: gap of unknown length  
\* 73918 74017: contig of 1385 bp in length  
\* 74017 77137: gap of unknown length  
\* 77138 77237: contig of 3120 bp in length  
\* 77238 79997: gap of unknown length  
\* 79998 80097: contig of 2760 bp in length  
\* 80098 82859: gap of unknown length  
\* 82859 82859: contig of 2762 bp in length  
\* 82860 85680: gap of unknown length  
\* 85681 85780: contig of 2721 bp in length  
\* 85781 89551: gap of unknown length  
\* 89551: contig of 3771 bp in length.





```

repeat_region /rpl_family="L1"
10850..11026 /rpl_family="L2"
repeat_region 11461..117576
/rpl_family="L1"
repeat_region 17680..17758
/rpl_family="Mariner"
repeat_region 18412..18629
/rpl_family="MIR"
repeat_region 19942..20121
/rpl_family="MER1_type"
repeat_region 20456..20629
/rpl_family="(TA)n"
repeat_region 20645..20757
/rpl_family="L1"
repeat_region 20758..21114
/rpl_family="MALR"
repeat_region 22734..23101
/rpl_family="MALR"
repeat_region 23102..24334
/rpl_family="L1"
repeat_region 24340..24720
/rpl_family="L1"
repeat_region 24721..25077
/rpl_family="L1"
repeat_region 25078..25115
/rpl_family="(TA)n"
repeat_region 25116..25599
/rpl_family="L1"
repeat_region 25600..25625
/rpl_family="AT_rich"
repeat_region 26058..26331
/rpl_family="L2"
repeat_region 26345..26475
/rpl_family="L2"
repeat_region 27007..27085
/rpl_family="GA_rich"
repeat_region 27113..27146
/rpl_family="(CA)n"
repeat_region 27479..27605
/rpl_family="MER1_type"
repeat_region 27924..28079
/rpl_family="MER1_type"
repeat_region 29818..30292
/rpl_family="MER1_type"
repeat_region 31387..31688
/rpl_family="Alu"
repeat_region 32916..32981
/rpl_family="CT_rich"
repeat_region 32982..33240
/rpl_family="CT_rich"
repeat_region 33241..33252
/rpl_family="Alu"
repeat_region 34110..34142
/rpl_family="AT_rich"
repeat_region 35992..36014
/rpl_family="(TAAA)n"
repeat_region 36034..36187
/rpl_family="Retroviral"
repeat_region 36188..36481
/rpl_family="Alu"
repeat_region 36482..36514
/rpl_family="Retroviral"
repeat_region 36617..36835
/rpl_family="L1"
repeat_region 36838..36905
/rpl_family="(TG)n"
repeat_region 36906..36939
/rpl_family="(TA)n"
repeat_region 36945..37725
/rpl_family="L1"
repeat_region 37736..38050
/rpl_family="Alu"

```

```

repeat_region 38070..38557
/rpl_family="MER4_group"
repeat_region 38558..38976
/rpl_family="L2"
repeat_region 39061..39116
/rpl_family="(TAAA)n"
repeat_region 39477..39693
/rpl_family="MIR"
repeat_region 40301..40339
/rpl_family="(TC)n"
repeat_region 40405..40630
/rpl_family="Alu"
repeat_region 40713..40816
/rpl_family="MER1_type"
misc.feature 41875..42095
/feature="match to EST AL046241 (NTID:95434325)"

Query Match 11.6% Score 147.6; DB 9; Length 138890;
Best Local Similarity 73.7%; Pred. No. 3.2e-20;
Matches 202; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

QY 831 CTGCAGAGACTGCAGAAATCCGGTGGAGAGACGCCCTATGTGTGCTGTGCGAGGCACAG 890
|||||
Db 28657 CTCGAGAGACTGTGACATCCAAATGGAGGGGTGCTTATGTGTGTATAGAGACTTAA 28598

QY 891 TAAGGTCTTACTCAACACATCTCTCGCGAGTGTGTGAGAAATCTTCACACCTTACAGAG 950
|||||
Db 28597 TAAGAACTACTTCAA--AGTTTACAGAACTGTGTGAGAGCTTTAACTCGTACCAG 28541

QY 951 TGGAGACATCACCTGCAAGTTGTAAGAAGACTGATGTGCTGCCCGACGACATTC 1010
|||||
Db 28540 GGGAGACATACAGCAATCAAGTTGTAATAAATGATGTGCTGCTAGCAAGATTC 28481

QY 1011 GCCAGTGTGACCTTAAGGCCCCATCGCAGACAGACTTGTGTGAGATGCAGACAAAC 1070
|||||
Db 28480 TCTACATAGACCTCAAGGCCCTATGTGCAAGATTTGTGTGAGATGCAGACAAAC 28421

QY 1071 GCGTGTCTCGCAGACAGACTTACGCTTCAATA 1104
|||||
Db 28420 ACCTATCTGTGTGAGACACTTACGCTTCAATA 28387

RESULT 11
AC037443 151961 bp DNA linear HTG 24-AUG-2002
LOCUS AC037443
DEFINITION Homo sapiens chromosome 2 clone RP11-77F17 map 2, WORKING DRAFT
ACCESSION AC037443
VERSION AC037443.2 GI:7960458
SOURCE HTG; HTGS_PHASE1; HTGS_DRAFT.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 151961)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 2, clone RP11-77F17
Unpublished
2 (bases 1 to 151961)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collins,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glende,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karakas,A.,
Klein,J., Lacombe,K., Lamazares,R., Lander,T., Lehotzky,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGuirk,A., McKernan,K., McPheters,R.,
Melarini,J., Menes,L., Mihova,T., Miranda,C., Mlenda,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

```

# REFERENCE JOURNAL

O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisanil, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Koltman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

# TITLE

Direct Submission  
Submitted (08-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
(bases 1 to 151961)

# AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abramson, H., Allen, N., Anderson, S., Baldwin, J., Barnum, N., Bastien, V., Beda, F., Boguslavsky, L., Bouckhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choquel, Y., Delange, M., Collins, S., Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J., S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J., Ilev, I., Johnson, R., Jones, C., Kamp, L., Karatas, A., Klein, O., Lacombe, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McGuire, A., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisanil, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

# COMMENT

Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 21, 2000 this sequence version replaced gi:7528172.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

# TITLE

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

# JOURNAL

Project Information  
Center project name: L59226  
Center clone name: 77\_F\_17

# REFERENCE

Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 147226 bases at least Q40  
Consensus quality: 149421 bases at least Q20  
Consensus quality: 150378 bases at least Q20  
Insert size: 194000; agarose-ff  
Insert size: 151061; sum-of-ctrls  
Quality coverage: 4.5 in Q20 bases; agarose-ff  
Quality coverage: 5.8 in Q20 bases; sum-of-ctrls

# AUTHORS

NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

# COMMENT

1 1451: contig of 1451 bp in length  
\* 1452 1551: gap of 100 bp  
\* 1552 3695: contig of 2144 bp in length  
\* 3696 3795: gap of 100 bp  
\* 3796 6752: contig of 2957 bp in length  
\* 6753 6852: gap of 100 bp

FEATURES  
source  
1. 151961  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2"  
/clone="RP11-77F17"  
/clone\_1bp="RP11-77F17 Human Male BAC"  
1. 1451  
/note="assembly-fragment"  
1552. 3695  
/note="assembly-fragment"  
3796. 6752  
/note="assembly-fragment"  
clone\_end:77  
vector\_side:right"  
6853. 18026  
/note="assembly-fragment"  
18127. 30673  
/note="assembly-fragment"  
30774. 44941  
/note="assembly-fragment"  
45042. 58179  
/note="assembly-fragment"  
clone\_end:SP6  
vector\_side:left"  
58280. 74554  
/note="assembly-fragment"  
74555. 105703  
/note="assembly-fragment"  
105804. 151961  
/note="assembly-fragment"

BASE COUNT 47790 a 28505 c 28546 g 46219 t 901 others  
ORIGIN

Query Match 11.6%; Score 147.6; DB 2; Length 151961;  
Best Local Similarity 73.7%; Pred. No. 3.2e-20;  
Matches 202; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

831 CTGCAAGACGCAAAATCCGGTGGAGAGCGCCTATGCTGCTGCGAGGCGCACCAG 890  
|||||  
41665 CTGCAAGACGCTGTCATCCATGAGGAGGCTGCTATGCTGCTGCTATGAGACCTAA 41606  
|||||  
891 TAAAGTGTACTTCAAAAGCTTGCAGTGTGGAGAAATCTTCAACCTTACAGAG 950  
|||||  
41605 TAAAGTGTACTTCAAAAGCTTGCAGTGTGGAGAAATCTTCAACCTTACAGAG 41549  
|||||  
951 TGGAGACATCATCCTGCAAAAGTTGTAAGAACTAGATGCTGCCCTGACCTGATTTTC 1010  
|||||  
41548 GGGAGATTAACAGCAATCAAGTTGTAATTAAGTATGCTGCTGCTGAGCAAGATTTTC 41489  
|||||  
1011 GCCAGTGGACCTTAAAGCCCATCGGCAAGACTTGTGGAGAGTCAAGAGCAAAAC 1070  
|||||  
41488 TCTACATAGACCTTAAAGCCCATTTGTCAGATTTGTGGAGAGTCAAGAGCAAAAC 41429  
|||||  
1071 GCCTGGCCGAGCAGCACTTACCTTCAATA 1104  
|||||  
41428 ACCTATCCTGTGTGAGCACTTTCAGCTTCAATA 41395  
|||||



|           |  |
|-----------|--|
| REFERENCE | 2 (bases 1 to 187578)  |
| AUTHORS   | Mang,C., Meyer,R. and Dignan,G.  |
| TITLE     | The Sequence of Homo sapiens BAC clone RP11-191J2  |
| JOURNAL   | Unpublished (2001)   |
| REFERENCE | 3 (bases 1 to 187578)  |
| AUTHORS   | Waterston,R.H.   |
| TITLE     | Direct Submission  |
| JOURNAL   | Submitted (03-OCT-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA |
| REFERENCE | 4 (bases 1 to 187578)  |
| AUTHORS   | Waterston,R.H.   |
| TITLE     | Direct Submission  |
| JOURNAL   | Submitted (29-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA |
| REFERENCE | 5 (bases 1 to 187578)  |
| AUTHORS   | Waterston,R.H.   |
| TITLE     | Direct Submission  |
| JOURNAL   | Submitted (05-APR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA |
| REFERENCE | 6 (bases 1 to 187578)  |
| AUTHORS   | Waterston,R.   |
| TITLE     | Direct Submission  |
| JOURNAL   | Submitted (20-APR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA                 |
| COMMENT   | On Apr 5, 2002 this sequence version replaced gi:19807902.   |

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

**SOURCE INFORMATION:** The RpCI-11 human BAC library was made from the blood of one male donor, as described by Goegeawa, K., Moon, P. Y., Zhao, B., Frengen, E., Teleno, M., Cattnese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

**NEIGHBORING SEQUENCE INFORMATION:** The clone sequenced to the left is RP11-473N2; the clone sequenced to the right is RP11-575K2, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-19J2; actual end is at base position 17325 of RP11-755K2.

Polymorphisms have been identified between AC032000, AC024191, and

```
AC096652. Data from AC032000 and AC024191 were used to finish this
clone, AC096952.
Location/Qualifiers
1..187578
```

|               |  |
|---------------|--|
| misc_feature  | 1..151<br>/note="match to EST R34463 (NID:g791364) yh85c01.r1"                       |
| repeat_region | 659..698<br>/rpt_family="L2"   |
| repeat_region | 1017..1100<br>/rpt_family="L1"   |
| misc_feature  | 4161..4181<br>/note="match to EST BI496001 (NID:g153535345)"                         |
| repeat_region | 4163..4477<br>/rpt_family="Alu"  |
| misc_feature  | 4178..4188<br>/note="match to EST AA749271 (NID:g27892229) ny04d05.s1"               |
| repeat_region | 4541..4567<br>/rpt_family="TTTG)n"   |
| misc_feature  | 4557..4577<br>/note="similar to Homo sapiens EST AI265985 (NID:g3884143) qm46c01.x1" |
| misc_feature  | 5981..5992<br>/note="match to EST BI493951 (NID:g15333295)"                          |
| repeat_region | 8924..9119<br>/rpt_family="L1"   |
| repeat_region | 9141..10080<br>/rpt_family="L1"  |
| repeat_region | 10292..10618<br>/rpt_family="L1"   |
| misc_feature  | 10536..11223<br>/note="match to EST BG506934 (NID:g13468451)"                        |
| repeat_region | 10662..10970<br>/rpt_family="L1"   |
| repeat_region | 11491..11819<br>/rpt_family="L1"   |
| repeat_region | 11832..12135<br>/rpt_family="Alu"  |
| repeat_region | 12137..12205<br>/rpt_family="MER2_type"  |
| repeat_region | 12204..12819<br>/rpt_family="MER2_type"  |
| repeat_region | 12830..13102<br>/rpt_family="L1"   |
| repeat_region | 13109..13255<br>/rpt_family="MER2_type"  |
| repeat_region | 13256..13378<br>/rpt_family="L1"   |
| repeat_region | 13379..13664<br>/rpt_family="Alu"  |
| repeat_region | 13665..14289<br>/rpt_family="L1"   |
| misc_feature  | 14301..14515<br>/note="match to EST BG995513 (NID:g14399583)"                        |
| misc_feature  | 14357..14790<br>/note="similar to Homo sapiens EST BE735137 (NID:g10194129)"         |
| misc_feature  | 14405..14880<br>/note="similar to Mus musculus EST BI407638 (NID:g15168561)"         |
| misc_feature  | 14409..14845<br>/note="similar to EST BM912100 (NID:g19362479)"                      |
| misc_feature  | 14441..14845<br>/note="similar to EST B0054284 (NID:g19813624)"                      |
| misc_feature  | 14447..14827<br>/note="similar to Homo sapiens EST BE799707 (NID:g10220905)"         |
| misc_feature  | 14447..14813   |



FEATURES Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA  
 Location/Qualifiers  
 source 1.125020  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="16"  
 /map="16q24.3: between D16S520 and WI-12410"  
 /note="1 isolated from a patient with Huntington's disease like 2 (HDL2)"  
 repeat\_region complement(3581..35746)  
 /rpt\_type="tandem"  
 /rpt\_unit="ctg"  
 complement(<36507..>36887)  
 gene /gene="JPH3"  
 /note="Jp3"  
 complement(<36507..>36887)  
 mRNA /gene="JPH3"  
 /product="junctional protein 3"  
 complement(<36507..36887)  
 CDS /gene="JPH3"  
 /note="component of the junctional complex between plasma membrane and endoplasmic reticulum"  
 /codon\_start=1  
 /product="junctional protein 3"  
 /protein\_id="AAL40941.1"  
 /db\_xref="GI:17646245"  
 /translation="MSSSGRPNFDDGGSCGCGWEDGKAHGHGCTGKGGGEGYTSMS  
 HGFVLYGWTPSGNTYGTGTAQGRHIGLESKGVYKGEWTHGFRGRTYRECAG  
 NGAKEGTWMSGLDQGYETETYSdg"

BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others  
 ORIGIN

Query Match 5.7% Score 72.2; DB 9; Length 125020;  
 Best Local Similarity 11.1%; Pred. No. 0.0011;  
 Matches 100; Conservative 401; Mismatches 395; Indels 4; Gaps 3;

QY 4 GCGGGGAGCGGCGGAGCGACCATGTCCTCCGCGAGCAGCGTCCACCCCTGCCGCAT 63  
 Db 17783 GSKSSGCGTGYIYKKKGKGGKSSKMKWGSSTRSSSAKSCSYMGMCMGMSMA 17724  
 QY 64 CCTTATCCGAGCGCACCAAGCGGGGATGGCTGGAGGTTCGAGACCGAGGGCTGCGA 123  
 Db 17723 KSYMICYMYRMSYSYKYCYSCMGMSSTYSYSCMKMSWGSCKYKMKYYSGWSS 17664  
 QY 124 CCGCGCGCCCTCTCTCTCCCGGCTACAGACAGCTCATGCGCGCGGAGTACGCA - 182  
 Db 17663 YSTSMGYSTSTCKKYKCSMSMYKCTSKYRKRSTYYWGGRKAKKYYCAGHRRM 17604  
 QY 183 -CAGCCACGAGCGGCGACAGCTCATGGCCCTGCTGCGGATGGGTCCCGGTCGTC 241  
 Db 17603 SYMKCCAKMMMSYCMWYYSKCTYKSSCTYKRGYWGSKTYKSAGKSRSMY 17544  
 QY 242 GCAGCCGTGACGCTGCGTGACG-GTGAACCGCGCGCGAGCGCTCGGTGCACTTCA 300  
 Db 17543 CMRRSSKSSSMMSMASSMCMGAGYRKRSAAGWRSKGRSTGKMRACSKRT 17484  
 QY 301 CTGGGGCGCGCGAGCGCTGAGGCTGCGAGCGACCGCCCGCCGATCGGCT 360  
 Db 17483 GSYGSTRSRKKKGYSKYRKGKKTCTYCMWYKRYKRTSMCWYTKMSGYRYRKR 17424  
 QY 361 TCCTGTCAACCCCGTGGCGACCGCGCGGAGGATCCCGCG-ATCCTGCGAGACGT 419  
 Db 17423 CCKKKKCTGYRMSKSKSYKSMRGRSSYSTCMWSOMGYSMKCMKMYSKKRRS 17364  
 QY 420 AGCCCGCTCTGCTGCTGACCTTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 479  
 Db 17363 MRGSMSSKMRGYAGRCYSSSMWTRKRKRSKCYSYKYGKMKMGWGMKRGSKYWS 17304  
 QY 480 CAGGCGACACCGACGAGGAGGAGGAGCGCGGATCCTCGGAGCGGAGACCGGA 539  
 Db 17303 SMKMKRSSWSKCTYKSYSGRRKSGWGRSTKRAKSSMRMAGSKCTYGSSTYNNRR 17244  
 QY 540 GCCGAGAGAGTGGCGCGAGGAAAGCGTCCCGACCGCGAGGAGGAGGCGATGT 599

Db 17243 NNRMGKTGCVNMYRRSRAMNNGNAAAAGCTTCCCCANTNGCGGAAAAGGCGGSASRAS 17184  
 QY 600 TCAGGCTGACAGGCGAGCGCGCGGTGGAGACGACACCGACCGAGAGACCGAACAAGTGN 659  
 Db 17183 CYKGRMSSKSCYRSGTRRCMKSKRCRSGSKSMGMYTRGNSGKTSYSAKSGRCYTCW 17124  
 QY 660 GCGCGCATCGAGTCTAGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 719  
 Db 17123 GWRGKGRKCMSSRKMKYKSYRRKRMTCMKMYSYMMYRSMCMCMKSCCGCY 17064  
 QY 720 GGACCCCGGTGATTCGATGCCCCCTGAGACACGAGCTCCCCGCAAGCAGGACAGA 779  
 Db 17063 MGMSYSYSGKYSWGMKSYMWRYSKSTSKAWRSSKRGMTGGRYKGGRSYGRGGGS 17004  
 QY 780 CAAGACGCGCTGGCTTCCAGTCTTAGACGAGCAAGTACGGCTACATCAGTCAAGA 839  
 Db 17003 YKGGGSMGKGTGAGGAGAGCGGCTATGTTGTGTGTGTCAGGCGACAGTAAGTGT 899  
 QY 840 CTGCAAAATCCGCTGAGGAGAGCGGCTATGTTGTGTGTGTCAGGCGACAGTAAGTGT 899  
 Db 16943 CWSSCWMSMKKSKGSMKAGWWDKGSYSTDKSDKMBRSBSKYKSKWMSGRGYY 16884

Search completed: January 19, 2003, 22:13:21  
 Job time : 5277 secs



IS-07-745-206A-12

Query Match 4.2% Score 54; DB 1; Length 5467;  
 Best Local Similarity 46.7%; Pred. No. 0.00048;  
 Matches 171; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```

303 CGGGGCGCCGACGCTGCACGCTTCGCGAGCCGCGGAGCGCCGCGATCGGGTTTC 362
    ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2800 CGCGCGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2859
    ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 363 CTGTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCGCGATCTTGGACACCGT 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2860 CGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 CCGGTTCTGCTGCTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCT 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2920 GCGCGCGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 483 GCGAGACACCCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2980 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 543 GAGAGAGGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3040 CGCGGAGCGGAGAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3099
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 603 GCGTCGAGGAGCGCGCGGTTGGAGACGACGACCGCGGAGGAGGAGGAGG 662
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3100 GCGTCGTCACGAGGCTGTGTGAGAGGAGACACGAGGAGGAGGAGGAGGAG 3159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 663 GCGCAT 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3160 TGAGAT 3165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 2  
 US-08-311-363-12  
 ; Sequence 12, Application US/08311363  
 ; Patent No. 5876958

GENERAL INFORMATION:  
 APPLICANT: Harpold, Michael  
 APPLICANT: Ellis, Steven  
 APPLICANT: Williams, Mark  
 APPLICANT: Feldman, Daniel  
 APPLICANT: McCue, Ann  
 TITLE OF INVENTION: Human Calcium Channel Compositions and  
 TITLE OF INVENTION: Methods  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Brown, Martin, Haller & McClain  
 STREET: 1660 Union Street  
 CITY: San Diego  
 STATE: California  
 COUNTRY: USA  
 ZIP: 92101-2926  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/311,363  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/745,206  
 FILING DATE: 15-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Seidman, Stephanie L.  
 REGISTRATION NUMBER: 33,779  
 REFERENCE/DOCKET NUMBER: 6362-51506  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619)238-0999  
 TELEFAX: (619)238-0062  
 INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 5467 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390..3392, 3396..3488, 3495..3539, 3543..3581, 3585..3587, 3591..3626, 3630..3689, 3693..3737, 3744..3746, 3750..4823, 4827..4841, 4845..5006, 5010..5096, 5100..5306, 5310..5366, 5370..5465)

Query Match 4.2% Score 54; DB 2; Length 5467;  
 Best Local Similarity 46.7%; Pred. No. 0.00048;  
 Matches 171; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```

303 CGGGGCGCGACGCTGCAGCGGTGCGGAGCCGCGGAGCGCCGCGATCGGGTTTC 362
    ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2800 CGCGCGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2859
    ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 363 CTGTCAACCCCGTGGCCACGCGCGCGGAGATCCCGCGATCTTGGACACCGT 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2860 CGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 CCGGTTCTGCTGCTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCT 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2920 GCGCGCGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 483 GCGAGACACCCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2980 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 543 GAGAGAGGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3040 CGCGGAGCGGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3099
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 603 GCGTCGAGGAGCGCGCGGTTGGAGACGACGACCGCGGAGGAGGAGGAGG 662
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3100 GCGTCGTCACGAGGCTGTGTGAGAGGAGACACGAGGAGGAGGAGGAGGAG 3159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 663 GCGCAT 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3160 TGAGAT 3165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3  
 US-08-232-463-14/c  
 ; Sequence 14, Application US/08232463  
 ; Patent No. 5670367

GENERAL INFORMATION:  
 APPLICANT: DORNER, F.  
 APPLICANT: SCHEFFLINGER, F.  
 APPLICANT: FALKNER, F. G.  
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 1800 Diagonal Road, Suite 500  
 CITY: Alexandria  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22313-0299  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/232,463  
 FILING DATE:

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)83-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZqpt-Fls
US-08-232-463-14

```

```

Query Match 4.2%, Score 53.2, DB 1: Length 7218;
Best Local Similarity 6.5%; Pred. No. 0.00085;
Matches 25; Conservative 202; Mismatches 155; Indels 0; Gaps 0;

```

```

OY 465 GGAGTTCGGGAGCGAGACACCCAGAGAGAGGAGCGCCGCGCATCTCGGG 524
DB 1388 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1329
OY 525 GACCCGGAGACCGGAGCGGAGAGGTGCGCCGCGAAGCGCCCGCCGCGAG 584
DB 1328 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1269
OY 585 CGAGAGCGCCATGTCAGGCTGCAGCGCAGCGCGGTGGAGAGACACCGAGA 644
DB 1268 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1209
OY 645 GGACCGAAGAGTGGCGGCGATGTCAGCTGGAGCGAGGAGCGCATCTCGTC 704
DB 1208 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1149
OY 705 CGCAGAGTGGCTCAGACCGCGGTGATTCGATGCCCTGAGACCGAGCTCCCGCA 764
DB 1148 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1089
OY 765 AAGCAGGAGAGAGAGAGAGCGCGCGCTGCTTCCAGTCTTGAAGAGAGTACGGCTA 824
DB 1088 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1029
OY 825 CTATCAGTGAAGAGCTGCAAA 846
DB 1028 AATCTGTGAGCGTATGCAAA 1007

```

```

RESULT 4
US-08-456-200B-11
Sequence 11, Application US/08456200B
Patent No. 6229000
GENERAL INFORMATION:
APPLICANT: Franz, Jurgen; Weingartner, Bernhard;
APPLICANT: Unterbeck, Axel; Rae, Peter
TITLE OF INVENTION: TISSUE-SPECIFIC HUMAN NEURONAL
TITLE OF INVENTION: CALCIUM CHANNEL SUB-TYPES AND
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SPRUNG HORN KRAMER & WOODS
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York

```

```

COUNTRY: U.S.A.
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: NEC Powermate SX/20
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,200B
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/094,712
FILING DATE: 19-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/858,278
FILING DATE: 26-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/064,778
FILING DATE: 19-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 41 10 785
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Bayer 8398.3-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 6232 nucleotides
TYPE: Nucleotide
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-456-200B-11

```

```

Query Match 4.0%; Score 50.8; DB 4: Length 6232;
Best Local Similarity 48.0%; Pred. No. 0.0032;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

```

```

OY 250 GAGCTGCGGTGAGTGAACCCGCGCGAGCGCTCGTGCAAGTTCACGCGCGC 309
DB 1636 GAGGCCCGGAGAGCGGAGAGCGGAGCGCGGCGCGGAGAGAGCGCGCGCGC 1695
OY 310 CGCAGCTGAGCTGCGAGGTCGAGGTCGAGGTCGAGGTCGAGGTCGAGGTC 369
DB 1696 CGCA-GCCACAGCAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 1752
OY 370 CCCGTCGCGCAGCGCGCGCGGAGATCCCGCGATCTCTGACAGACCTGACCGCTTC 429
DB 1753 GCGCCAGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1812
OY 430 TCGTCCGTGACCTTCTGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 1813 GAGCGGAGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1872
OY 490 CCCACGAGAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 549
DB 1873 GGGGCGCAAGGCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1932
OY 550 GTGGCGCGAGAGAAAGCGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
DB 1933 GCGGAGAGCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1992
OY 610 GGGCAGCGCGGTCGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 669
DB 1993 CACGAGGCTGTGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2051

```



| Matches | 201: Conservative  | 0: Mismatches | 215: Indels | 3: Gaps |
|---------|--|---------------|-------------|---------|
| QY      | 250 GACGGTGGCGGTCCAGGTGAACCGCGCGCGACAGCTCTCGTGCAGTGTCTACTCGGGCCC | 309           |             |         |
| Db      | 2751 GAGGCCCCCAAGAGCGGAGCGGGAGCCCGGTGCCCGGAGAGCGCGCGCCGAC        | 2810          |             |         |
| QY      | 310 CGCAGCTTCACCCCTGCAGGTGTCCGAGCCGACCCCGACGCCGATCCGGTTCCTGTCA   | 369           |             |         |
| Db      | 2811 CGCA--CGCCACGACGAGAGCGCGGGGGCC--CCGGAGAGCGCGAGCGAGCGCGGCCA  | 2867          |             |         |
| QY      | 370 CCCCCTGGCCAGCGCGCGCGCGGAGATCCCGCGATCTCTGCGACGCTAGCCCCGTTC    | 429           |             |         |
| Db      | 2868 GGCACAGGCGCCCGAGCGCGCGCGCGGCGGACACACCGCGCGCTCCCGGAGAGCGCGCC | 2927          |             |         |
| QY      | 430 TGTCTCGTACCTTCTGTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  | 489           |             |         |
| Db      | 2928 GAGCGGGAGACCCCGACGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG     | 2987          |             |         |
| QY      | 490 CCCACGAAGGAGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG    | 549           |             |         |
| Db      | 2988 GCGCGCAAGAGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG     | 3047          |             |         |
| QY      | 550 GTGGCGCGGAGGAAAGCGGTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG     | 609           |             |         |
| Db      | 3048 GCGGAGAGCGGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG      | 3107          |             |         |
| QY      | 610 GGGCGAGCGCGGTGGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG       | 668           |             |         |
| Db      | 3108 CACGAGGCTGTGTGAGAGAGAGACGCGGAGAGAGGCGCGGAGAGAGAGGCTGAGAT    | 3166          |             |         |

```

TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-193-078B-8

```

```

Query Match          3.8% Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0094;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

```

```

QY 250 GACGCTGCGTGTCAGTGAACCCGCGCGCAGCCGCTGGTGCACTGTTCACTGGGCGC 309
DB 2751 GAGGCCCCGAAGAGCGGAGAGCGGAGCCCGGTGCGCGGAGAGAGCGCGCGCGCAC 2810
QY 310 CGCAGCGTGCAGCTGCAGAGGTCGCGAGCCGCGACGCCCGCATCGGGTTCCTGTCAA 369
DB 2811 CGCA-GCCACACAGAGAGCGCGCGGCGC--CCGAGAGCGCGGAGAGCGCGCGCGCA 2867
QY 370 CCGCGTGCGCAGCGCGCGCGCGGAGATCCCGCGCATCTGGCAGACCGCTGACCGCCGTT 429
DB 2868 GCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2987
QY 490 CCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549
DB 2988 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3047
QY 550 GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
DB 3048 GCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3107
QY 610 GGGCAGCGCGCGGTGGAGCAGCAGCCACCGCGAGAGAGCGGAGACAGTGTGGCGCAT 668
DB 3108 CACGAGGCTGTGAGAGAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGTGAT 3166

```

```

RESULT 9
US-08-223-305C-8
Sequence 8, Application US/08223305C
Patent No. 5851824
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 4, 1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0062
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-223-305C-8

```

```

Query Match          3.8% Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0094;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

```

```

QY 250 GACGCTGCGTGTCAGTGAACCCGCGCGCAGCCGCTGGTGCACTGTTCACTGGGCGC 309
DB 2751 GAGGCCCCGAAGAGCGGAGAGCGGAGCCCGGTGCGCGGAGAGAGCGCGCGCGCAC 2810
QY 310 CGCAGCGTGCAGCTGCAGAGGTCGCGAGCCGCGACGCCCGCATCGGGTTCCTGTCAA 369
DB 2811 CGCA-GCCACAGAGAGAGCGCGCGGCGC--CCGAGAGCGCGGAGAGCGCGCGCGCA 2867
QY 370 CCGCGTGCGCAGCGCGCGCGCGGAGATCCCGCGCATCTGGCAGACCGCTGACCGCCGTT 429
DB 2868 GCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2987

```

```
0y      490   CCCACGAGGGAGAGGGGAGCCCGGCCTCCTCGGGAGACC6GGGAACCCGAGCCGAGAGG    549  
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db      298   GCGGCCAAGAAGGGCGAGCCGGCCCCCACCGGGCCCCCCCAGCGGGCCCGGAG        304  
  
0y      550   GTGGCCCGGAGAAAGCGGTTCCTCCCAAGCGAGAGGAGCGCATGTTTCAAGCTGCA       609  
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db      3048  GCGGAGAGCGGGGAGAGAGCCGGCCGGCCGACC GGCCCGGCGACAAGCGCGACTGT       3107  
  
0y      610   GCGCAGCGCCGGTGTAAGCAGCAGCCACCAAGCGAGAGCGGAACAGTGTGGCGGCTG     658  
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db      3108  CACGAGGCTGTGGAAGAGAGACCACGAGAGAGAGGCCACGGAAGAAGAGGOTGAT       3166
```

RESULT 10  
US-08-149-097D-8  
Sequence 8, Application US/08149097D  
Patent No. 5874236  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
City: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READERABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,097D  
FILING DATE: 05-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/105,536  
FILING DATE: 11-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US92/06903  
FILING DATE: 14-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/914,231  
FILING DATE: 13-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/868,354  
FILING DATE: 10-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/482,384  
FILING DATE: 20-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,751  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/176,899

```

1      FILING DATE: 04-APR-1988
2      ATTORNEY/AGENT INFORMATION:
3      NAME: Seidman, Stephanie L
4      REGISTRATION NUMBER: 35,777
5      REFERENCE/DOCKET NUMBER: 6
6      TELECOMMUNICATION INFORMATION:
7      TELEPHONE: (619) 238-0999
8      TELEFAX: (619) 238-0062
9      INFORMATION FOR SEQ ID NO: 8:
10     SEQUENCE CHARACTERISTICS:
11     LENGTH: 7175 base pairs
12     TYPE: nucleic acid
13     STRANDEDNESS: double
14     TOPOLOGY: linear
15     MOLECULE TYPE: DNA (genomic)
16     FEATURE:
17     NAME/KEY: CDS
18     LOCATION: 144..6857
19     FEATURE:
20     NAME/KEY: 5'UTR
21     LOCATION: 1..143
22     FEATURE:
23     NAME/KEY: 3'UTR
24     LOCATION: 6855..7175
25     OS-08-149-097D-8

```

|                       |              |                   |                 |              |
|-----------------------|--------------|-------------------|-----------------|--------------|
| Query Match           | 3.88;        | Score 49;         | DB 2;           | Length 7175; |
| Best Local Similarity | 48.08;       | Pred. No. 0.0094; |                 |              |
| Matches 201;          | Conservative | 0;                | Mismatches 215; | Indels 3;    |
|                       |              |                   |                 | Gaps 2       |

|    |      |   |      |
|----|------|---|------|
| QY | 250  | GAGCGTCGGTGCAGGTGAAACCCGCCGCCGGCAGCGCTTCGTGGACATTTCACTCGGGCGC | 309  |
| Db | 2751 | GAGGCCCGGAAGGGGAGACGGGBAACCCTGGTCCCGGAGAAGCGGCCGCCGCCAC       | 281  |
| QY | 310  | CGCACGCTGCACGCTTCAGGAGTGCCAGCGACGCCGCCGATTCGGATTCTGTCAA       | 369  |
| Db | 2811 | CGCA-GCCACACGAAAGAGAGCCCGGGGGCC--CCCGAGCGCGGAGCGAGCGCGCCGA    | 286  |
| QY | 370  | CCCCGTGGCACCGCGCGCGCGGAGATGCCCGGATCTTGGAGACAGACCGTAGGCCCGTTTC | 429  |
| Db | 2868 | GGCCCAAGCCCCGAGGGCGGGCCGGCACACCGCGCGCTCCCGAGAGAGCGCGCC        | 292  |
| QY | 430  | TCTGTCTGACCTTTCTGTGGCTCTCTCTCTCTACTGAGAGTTGCGGAGGACGACGACA    | 489  |
| Db | 2928 | GAGCGGAGGCCGACCCACCCACCGCGCACCGGACACAGGATCGAGAGCAAGAGTGCGCC   | 298  |
| QY | 490  | CCCACGAAGGAGAGGAGGAGCCCGCATCTCTGGGGACCCGGGAACCGCGGACGAGAG     | 549  |
| Db | 2988 | GGCGCCAAGGGCGAGCGCGCGCGCGGCAGACCGGGCGGCCCCGACGGGGCCCGGGAG     | 304  |
| QY | 550  | GTGGCCCGCAGGAAGCGGTCCTCCCAACCGCGAAGCGAGAGGGCGATTTCAAGCTGCA    | 609  |
| Db | 3048 | GCGGAGAGCGGGGAGGAGCGCGCGCGCACCGGGCCCGGACAAAGCGACGCTTGCT       | 310  |
| QY | 610  | GGGCAGGCCGCGGTGGGAGCAGCACCAACCGCGAGACCGGACCGGAACGTGTGGCGGGAT  | 668  |
| Db | 3108 | CACGAGGCTGTGGAGAAGAGACGACGAGGAGAGGAGGCGACGAGAGAGAGGCTAGAT     | 3166 |

RESULT 11  
US-08-949-386-8  
Sequence 8, Application US/08949386  
Patent No. 6090623  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: McQue, Ann  
APPLICANT: Gillespie, Alison  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:

|    |                                |  |
|----|--------------------------------|--|
| 1  | ADDRESS:                       | Brown, Martin, Haller & McClain                                      |
| 2  | STREET:                        | 1660 Union Street  |
| 3  | CITY:                          | San Diego  |
| 4  | STATE:                         | California   |
| 5  | COUNTRY:                       | US   |
| 6  | ZIP:                           | 92101  |
| 7  | COMPUTER READABLE FORM:        |  |
| 8  | MEDIUM TYPE:                   | Floppy disk  |
| 9  | COMPUTER:                      | IBM PC compatible  |
| 10 | OPERATING SYSTEM:              | PC-DOS/MS-DOS  |
| 11 | SOFTWARE:                      | Patentlin Release #1.0, Version #1.25                                |
| 12 | CURRENT APPLICATION DATA:      |  |
| 13 | APPLICATION NUMBER:            | US/08/949,386  |
| 14 | FILING DATE:                   |  |
| 15 | CLASSIFICATION:                |  |
| 16 | PRIOR APPLICATION DATA:        |  |
| 17 | APPLICATION NUMBER:            | US/08/290,012  |
| 18 | FILING DATE:                   | 11-AUG-1994  |
| 19 | APPLICATION NUMBER:            | 08/149,097   |
| 20 | FILING DATE:                   | 5-NOV-1993   |
| 21 | PRIOR APPLICATION DATA:        |  |
| 22 | APPLICATION NUMBER:            | 08/105,536   |
| 23 | FILING DATE:                   | 11-AUG-1993  |
| 24 | ATTORNEY/AGENT INFORMATION:    |  |
| 25 | NAME:                          | Seldman, Stephanie L.  |
| 26 | REGISTRATION NUMBER:           | 33,779   |
| 27 | REFERENCE/DOCKET NUMBER:       | 519808   |
| 28 | TELECOMMUNICATION INFORMATION: |  |
| 29 | TELEPHONE:                     | (619) 238-0999   |
| 30 | TELEFAX:                       | (619) 238-0062   |
| 31 | INFORMATION FOR SEQ ID NO:     | 8:   |
| 32 | SEQUENCE CHARACTERISTICS:      |  |
| 33 | LENGTH:                        | 7175 base pairs  |
| 34 | TYPE:                          | nucleic acid   |
| 35 | STRANDEDNESS:                  | double   |
| 36 | TOPOLOGY:                      | linear   |
| 37 | MOLECULE TYPE:                 | DNA (genomic)  |
| 38 | FEATURE:                       |  |
| 39 | NAME/KEY:                      | CDS  |
| 40 | LOCATION:                      | 144..6857  |
| 41 | FEATURE:                       |  |
| 42 | NAME/KEY:                      | 5'UTR  |
| 43 | LOCATION:                      | 1..143   |
| 44 | FEATURE:                       |  |
| 45 | NAME/KEY:                      | 3'UTR  |
| 46 | LOCATION:                      | 6855..7175   |
| 47 | US-08-949-386-8                |  |
| 48 | Query Match                    | 3.8%; Score 49; DB 3; Length 7175;                                   |
| 49 | Best Local Similarity          | 48.0%; Pred. No. 0.0094;   |
| 50 | Matches 201; Conservative      | 0; Mismatches 215; Indels 3; Gaps 2                                  |
| 51 | OY                             | 250 GACGCTGGGTCAGGTGAACCCGCGCGCAGACGCCTCGTGTCACTGTTCATCGGGGCC 309    |
| 52 | DB                             | 2751 GAGGCCCGCCAAAGCGAAGCGGGAGCCCGGTGTCGCCGGAGAAGCGCGCGCCGAC 2810    |
| 53 | OY                             | 310 GCACGCTGCACCTGCAGAGTGTCGAGCGCACCCCGAGCCCGATCGGTTCTGTCAA 369      |
| 54 | DB                             | 2811 CGCA-GCCACACAGAGAGAGCGCGGGGCC--CCGGAGGCGCGAGCGAGCGCGGCCGA 2867  |
| 55 | OY                             | 370 CCCCGTGGCCACGCGGCGCGGAGATCCCCGCGATCTGTGCAGACCGTAGCCCCTTC 429     |
| 56 | DB                             | 2868 GCCCCAGCGCCCGAGGCGCGCGGCGCACACCGCGCGGCTCCCGGAGGAGGGGCC 2927     |
| 57 | OY                             | 430 TCGTCGTCACCTCTGTGGCGCTCTCTCATCTGAGAGTTGCGGAGGACAGGACAGACA 489    |
| 58 | DB                             | 2928 GAGCGGAGACCCCGACGCCACCGCGCACCGGACACCAAGATCCGACCAAGAGTCCGCC 2987 |
| 59 | OY                             | 490,CCACAGAGAGAGGAGGAGCCCGCATCTCGGGAGACCGGGAACCGGAGCCGAGAGAG 549     |
| 60 | DB                             | 2988 GGCGGCAAGGAGGAGGCGGCGGCGGACCGGCGGCGGCCCGCCGAGCGGGGCCCGGAG 3047  |
| 61 | OY                             | 550 GTGGCCGCGAGAAAGCGGTCCCGCACCGCGAAGCGAGAGGGCGATGTTCAGGCTGCA 609    |

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



FILED DATE: 14-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/745,206  
FILING DATE: 15-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/620,250  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/603,751  
FILING DATE: 08-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/482,384  
FILING DATE: 02-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US89/01408  
FILING DATE: 04-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/176,899  
FILING DATE: 04-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-519812  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 238-0999  
TELEFAX: (619) 238-0062  
INFORMATION FOR SEQ. ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..6857  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..143  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 6855..7175  
US-08-450-562-8

Query Match 3.8%; Score 49; DB 3; Length 7175;  
Best Local Similarity 48.0%; Pred. No. 0.0094;  
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGTGCAGTGAACCCGCGCGGACGCTCGGTGCACTGTTCACTCGGGCGC 309  
DB 2751 GAGGCCCGCGAAGGCGAGACGGGAGCCCGTCCCGGAGAGGCGCGCGCGCGCAC 2810  
QY 310 CGCAGCGTGCAGGCTGAGGGTGCAGCCAGCCCGGAGCCCGGATCGGGTTCTCTCAA 369  
DB 2811 CGCA-GCCACAGCAAGAGAGGCGCGCGGGCC--CCGGAAGCGCGGAGCGCGCGCGCA 2867  
QY 370 CCCGTTGGCCACCGCGCGCGGAGATCCCGGATCTCTGAGACCGTACCCCGTTTC 429  
DB 2868 GCGCCAGAGCGCGGAGCGCGCGCGCGCGCACACCGCGCGCTCCCGGAGAGGGCGCC 2927  
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGAGTTGCGGAGGACGACAGACA 489  
DB 2928 GACCGGAGCGCCCGACGCCACCGCGGACCGGACACGAGATCCGAGCAAGAGTCTCGCC 2987  
QY 490 CCCACGAGAGGAGAGGAGGAGCGCGGATCTCGGGGAGCGGAGAACCGGACCGGAGAGAG 549  
DB 2988 GCGGCCAGAGGCGAGCGCGCGCGCGGACCGCGGCGCGCGCGCGCGCGCGCGCGGAG 3047  
QY 550 GTGGCCGCGAGCAAGCGGTCCCGCAGCGCGAGAGCGGAGGCGGATTTTCAGGCTGCA 609  
DB 3048 GCGGAGAGCGGAGAGCGCGCGCGCGCGCGCACCGGCGCGCGCGCGCGCGCGCGCGCT 3107  
QY 610 GGGCAGGCGCGGTGGAGCAGCAGCACCAACCGGAGGAGCGGAGAACGTGTGGCGCGAT 668

DB 3108 CACGAGCGTGTGAGAGAGGAGACCCAGCAGAGAGAGGCGCGGAGAGAGGCTGACAT 3166  
RESULT 13  
US-08-984-709A-8  
Sequence 8, Application US/08984709A  
Patent No. 6320032  
GENERAL INFORMATION:  
APPLICANT: Williams, Mark E.  
APPLICANT: Stauderman, Kenneth A.  
APPLICANT: Harpold, Michael M.  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Heller Ehrman White & McAuliffe  
STREET: 4250 Executive Square, Suite 700  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/984,709A  
FILING DATE: 02-DEC-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 450-8400  
TELEFAX: (619) 587-5360  
INFORMATION FOR SEQ. ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..6857  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..143  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 6855..7175  
US-08-984-709A-8

Query Match 3.8%; Score 49; DB 4; Length 7175;  
Best Local Similarity 48.0%; Pred. No. 0.0094;  
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGTGCAGTGAACCCGCGCGGACGCTCGGTGCACTGTTCACTCGGGCGC 309  
DB 2751 GAGGCCCGCGAAGGCGAGAGCGGAGACCCGTTCCCGGAGAGGCGCGCGCGCGCAC 2810  
QY 310 CGCAGCTGCACTGCAAGGTGCGGACCGACCGCGGAGCGCGGATTCGTTCTCAA 369  
DB 2811 CGCA-GCCACAGCAAGAGAGGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGGAG 2867  
QY 370 CCCGTTGGCCACCGCGCGCGGAGATCCCGGATCTCTGAGACCGTACCCCGTTTC 429  
DB 2868 GCGCCAGAGCGCGGAGCGCGCGCGCGCGCACACCGGCGCGCGCTCCCGGAGAGGCGCC 2927  
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGAGTTGCGGAGGACGACAGACA 489

Db 2928 GAGCGGAGACCCCGACGCGCCGCGACCGCACACAGATCCAGCAAGAGAGTGCCTC 2987  
Qy 490 CCCAGAGAGGAGAGGAGCCCGCATCTTGGGGACCCGGAAACCGAGCCGAGAG 549  
Db 2988 GCGCGCAAGGAGGCGCGCGCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGAG 3047  
Qy 550 GTGGCCGCGGAGAAAGCGGTCCCGACCGCGAGCGAGAGGAGTTCAGGCTGCA 609  
Db 3048 GCGGAGAGCGGGGAGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 3107  
Qy 610 GCGGAGCGCGGTGGAGCAGCAGCCACCGAGAGAGCGGAGAGTGTGGCGGAT 668  
Db 3108 CACGAGCGTGTGAGAGAGACACGAGAGAGAGAGCGCGCGAGAGAGAGCTAGAT 3166

RESULT 14  
US-08-450-272-8  
Sequence 8, Application US/08450272  
Patent No. 6387696

GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: McCue, Ann  
APPLICANT: Gillespie, Allison  
APPLICANT: Feldman, Daniel  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: US  
ZIP: 92101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,272  
FILING DATE: 08/31,1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/404,950  
FILING DATE: 13-MAR-1995  
APPLICATION NUMBER: 08/336,257  
FILING DATE: 7-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/314,083  
FILING DATE: 28-SEPT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/311,363  
FILING DATE: 23-SEPT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/290,012  
FILING DATE: 11-AUG-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/223,305  
FILING DATE: 4-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/193,078  
FILING DATE: 07-FEB-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/149,097

FILING DATE: 5-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/105,536  
FILING DATE: 11-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/914,231  
FILING DATE: 13-JULY-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/868,354  
FILING DATE: 10-APR-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/06903  
FILING DATE: 14-AUG-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/745,206  
FILING DATE: 15-AUG-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-519812  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 238-0999  
TELEFAX: (619) 238-0062  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..6857  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..143  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 6855..7175  
US-08-450-272-8

Query Match 3.88; Score 49; DB 4; Length 7175;  
Best Local Similarity 48.08; Pred. No. 0.0094; Indels 3; Gaps 2;  
Matches 201; Conservative 0; Mismatches 215;

Qy 250 GACGTCGCGTGAGGTGAACCCGCGCGCGCGCTCGGTGAGTGTTCATCGGCGC 309  
Db 2751 GAGGCGCCGAGAGGCGGAGAGCGGCGGAGCGCGGTCTCCGCGGAGAGCGGCGCGCAC 2810  
Qy 310 GCGACGTCGACGCTGACGAGGTGCCGAGCCGAGCCGAGCGCGGATCGGCTCTGCAA 369  
Db 2811 CGCA-GCCACAGCAAGAGAGCGCGGCGGCGC--CCGAGAGGCGCGAGCGCGCGGA 2867  
Qy 370 CCCGCTGGCCACGCCGCGCGCGGAGATCCCCGGGATCTGGCGAGACCGAGCCGCTTC 429  
Db 2868 GCGCCAGGCGCCGAGGCGGCGCGCGCGGACCGCGCGCGGCTCTCCCGGAGAGCGCGCC 2927  
Qy 430 TCGTCGCGACCTTCTGTGGCTCTCTCTCACTGGAGTGGGAGGAGCGACAGCA 489  
Db 2928 GAGCGGAGGCCCGACGCGCGCGCGCGCGCGAGGAGATCCGAGCAAGAGTGGCC 2987  
Qy 490 CCCAGAGAGGAGAGGAGCGCGCATCTCTGGGGAACCGGAGACCGGAGAG 549  
Db 2988 GCGCGCAAGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047  
Qy 550 GTGGCCGCGGAGAAAGCGGTCCCGACCGCGGAAGCGAGAGGAGGCGATGTTCAGGCTGCA 609  
Db 3048 GCGGAGAGCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 3107

QY 610 GGGCAGGCGCGGTGGAGACGACCCAGCAGCCGAGACCGGAACTGTGGCGGCAT 668  
|||||  
Db 3108 CACGAGGCTGTGGAGAGAGACGACGAGAGAGGCGCACGAGAGAGGCTGTGAT 3166  
|||||

## RESULT 15

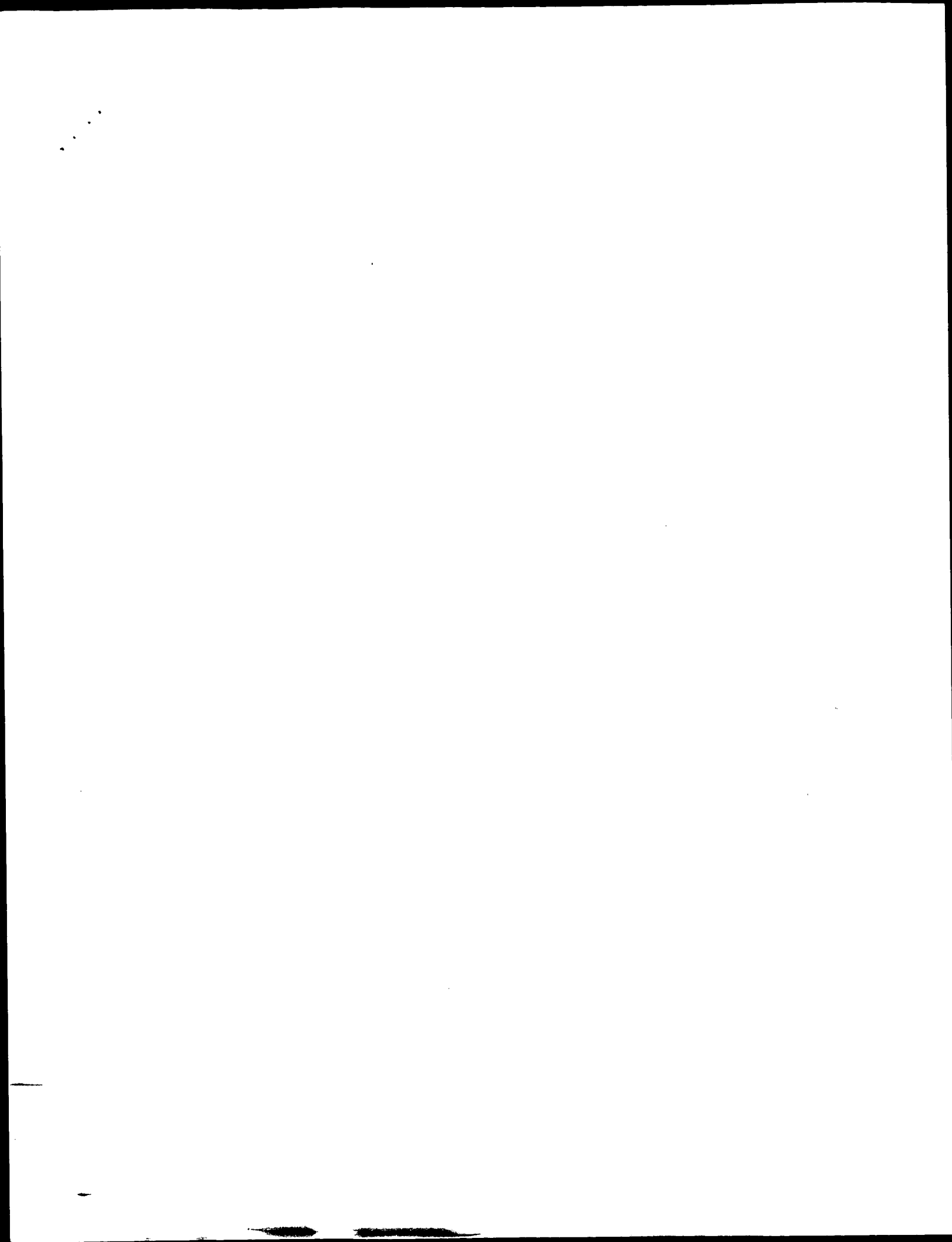
US-09-268-163-7  
; Sequence 7, Application US/09268163B  
; Patent No. 6353091  
; GENERAL INFORMATION:  
; APPLICANT: Lipscombe, Diane  
; APPLICANT: Schorge, Stephanie  
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF  
; FILE REFERENCE: B1055/7000  
; CURRENT APPLICATION NUMBER: US/09/268,163B  
; CURRENT FILING DATE: 1999-03-12  
; EARLIER APPLICATION NUMBER: US 60/077,901  
; EARLIER FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 7177  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 146..6856  
US-09-268-163-7

Query Match 3.8%; Score 49; DB 4; Length 7177;  
Best Local Similarity 48.0%; Pred. No. 0.0094;

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACCTGCGGTGACGAGTGAACCGCGCGCGCGCGCTGCGTGCAGTGTTCACCTCGGGCG 309  
|||  
Db 2753 GAGGCCCGCAAGCGGAGAGCGGAGCGCGGTGCGCGGAGAGAGCGCGCGCGCGCG 2812  
|||  
QY 310 CCGACGCTGACGCTGCAAGGCTGCGAGCCGACCGCGCGCGCGCGTTCCTGTCAA 369  
|||||  
Db 2813 CCGA-CGCAAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2869  
|||||  
QY 370 CCGCGTGGCGAGCG 429  
|||||  
Db 2870 GCGCGAGCG 2929  
|||||  
QY 430 TCGTCCGTGACCTTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489  
|||  
Db 2930 GAGCGGAGCG 2989  
|||  
QY 490 CCGACGAGGAGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 549  
|||||  
Db 2990 GCGCGCAAGGCG 3049  
|||||  
QY 550 GTGCGCGCGAGAGAGCGGTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609  
|||  
Db 3050 GCGGAGAGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3109  
|||  
QY 610 GGGCAGGCGCGGTGGAGACGACCCAGCAGCCGAGAGAGAGAGAGAGAGAGAGAG 668  
|||||  
Db 3110 CACGAGGCTGTGGAG 3168  
|||||

Search completed: January 19, 2003, 10:18:01  
Job time : 134 secs



GenCore version 5.1.3

1

(without alignments)  
9183.449 million cell updates/sec

US-09-844-864-1  
1077

Gapor 10.0 ,

787736

1

- 1: /cgn2\_6/prodata/1/pubnpa/us07\_PUBCOMB.seq \*
- 2: /cgn2\_6/prodata/1/pubnpa/ECT\_NEW\_PUB.seq \*
- 3: /cgn2\_6/prodata/1/pubnpa/ECT\_NEW\_PUB.seq \*
- 4: /cgn2\_6/prodata/1/pubnpa/us06\_NEW\_PUB.seq \*
- 5: /cgn2\_6/prodata/1/pubnpa/us07\_NEW\_PUB.seq \*
- 6: /cgn2\_6/prodata/1/pubnpa/ECTUS\_PUBCOMB.seq \*
- 7: /cgn2\_6/prodata/1/pubnpa/us08\_NEW\_PUB.seq \*
- 8: /cgn2\_6/prodata/1/pubnpa/us08\_PUBCOMB.seq \*
- 9: /cgn2\_6/prodata/1/pubnpa/us09\_NEW\_PUB.seq \*
- 10: /cgn2\_6/prodata/1/pubnpa/us09\_PUBCOMB.seq \*
- 11: /cgn2\_6/prodata/1/pubnpa/us10\_NEW\_PUB.seq \*
- 12: /cgn2\_6/prodata/1/pubnpa/us10\_PUBCOMB.seq \*
- 13: /cgn2\_6/prodata/1/pubnpa/us60\_NEW\_PUB.seq \*
- 14: /cgn2\_6/prodata/1/pubnpa/us60\_PUBCOMB.seq \*

**Pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

| Result No. | Score | Query Match | Length | DB | ID                 | Description        |
|------------|-------|-------------|--------|----|--------------------|--------------------|
| 1          | 1277  | 100.0       | 1277   | 10 | US-09-844-864-1    | Sequence 1, App1   |
| 2          | 799.4 | 67.6        | 814    | 10 | US-09-844-864-18   | Sequence 18, App   |
| 3          | 752.2 | 51.4        | 809    | 10 | US-09-844-864-22   | Sequence 22, App   |
| 4          | 273.6 | 21.4        | 305    | 10 | US-09-844-864-21   | Sequence 21, App   |
| 5          | 260.6 | 20.4        | 375    | 10 | US-09-844-864-25   | Sequence 25, App   |
| 6          | 98.8  | 7.7         | 123    | 10 | US-09-844-864-19   | Sequence 19, App   |
| 7          | 98.8  | 7.7         | 123    | 10 | US-09-844-864-23   | Sequence 23, App   |
| 8          | 72    | 5.6         | 105    | 10 | US-09-844-864-20   | Sequence 20, App   |
| 9          | 70.4  | 5.5         | 105    | 10 | US-09-844-864-24   | Sequence 24, App   |
| 10         | 49.6  | 3.9         | 14800  | 10 | US-09-954-456-1601 | Sequence 1601, App |
| 11         | 49    | 3.8         | 7177   | 12 | US-10-033-026-7    | Sequence 7, App1   |
| 12         | 49    | 3.8         | 7364   | 10 | US-09-954-456-1179 | Sequence 1179, App |
| 13         | 49    | 3.8         | 7364   | 12 | US-10-033-026-5    | Sequence 5, App1   |
| 14         | 49    | 3.8         | 7376   | 12 | US-10-033-026-3    | Sequence 3, App1   |
| 15         | 46.6  | 3.6         | 1006   | 9  | US-09-946-807-104  | Sequence 104, App  |
| 16         | 46.6  | 3.6         | 1006   | 10 | US-09-795-668-104  | Sequence 104, App  |
| 17         | 44.2  | 3.5         | 2481   | 10 | US-09-795-668-104  | Sequence 104, App  |
| 18         | 44.2  | 3.5         | 2481   | 10 | US-09-894-998-35   | Sequence 35, App   |
| 19         | 44.2  | 3.5         | 15872  | 9  | US-09-860-846-1    | Sequence 1, App1   |

|    |      |     |       |    |                     |                      |
|----|------|-----|-------|----|---------------------|----------------------|
| 21 | 44.2 | 3.5 | 15870 | 10 | US-09-861-289-1     | Sequence 1, Appl     |
| 20 | 44   | 3.4 | 14802 | 10 | US-09-954-456-1601  | Sequence 1601, Appl  |
| 22 | 43.4 | 3.4 | 2061  | 10 | US-09-962-832-239   | Sequence 239, Appl   |
| 23 | 43.4 | 3.4 | 2061  | 10 | US-09-954-456-1593  | Sequence 1593, Appl  |
| 24 | 42.4 | 3.3 | 4257  | 9  | US-09-825-288A-1    | Sequence 1, Appl     |
| 25 | 42.2 | 3.3 | 3203  | 10 | US-09-880-107-2164  | Sequence 2164, Appl  |
| 26 | 41.4 | 3.2 | 2724  | 12 | US-10-044-090-309   | Sequence 309, Appl   |
| 27 | 41.4 | 3.2 | 68750 | 9  | US-10-014-717-1     | Sequence 1, Appl     |
| 28 | 41.2 | 3.2 | 562   | 10 | US-09-925-301-525   | Sequence 525, Appl   |
| 29 | 41.2 | 3.2 | 2389  | 10 | US-09-799-875-7     | Sequence 7, Appl     |
| 30 | 41   | 3.2 | 4936  | 10 | US-09-967-768A-116  | Sequence 116, Appl   |
| 31 | 40.8 | 3.2 | 2561  | 9  | US-09-976-740-48    | Sequence 48, Appl    |
| 32 | 40.8 | 3.2 | 2561  | 9  | US-09-976-740-48    | Sequence 48, Appl    |
| 33 | 40.8 | 3.2 | 2561  | 12 | US-10-023-529-48    | Sequence 48, Appl    |
| 34 | 40.8 | 3.2 | 2561  | 12 | US-10-023-529-48    | Sequence 48, Appl    |
| 35 | 40.8 | 3.2 | 2561  | 12 | US-10-023-523-48    | Sequence 48, Appl    |
| 36 | 40.8 | 3.2 | 2561  | 12 | US-10-023-523-48    | Sequence 48, Appl    |
| 37 | 40.8 | 3.2 | 3233  | 10 | US-09-964-824-116   | Sequence 116, Appl   |
| 38 | 40.8 | 3.2 | 4257  | 9  | US-09-825-288A-1    | Sequence 1, Appl     |
| 39 | 40.4 | 3.2 | 804   | 10 | US-09-749-278B-62   | Sequence 62, Appl    |
| 40 | 40.4 | 3.2 | 1320  | 10 | US-09-921-398-28    | Sequence 28, Appl    |
| 41 | 40.4 | 3.2 | 2137  | 10 | US-09-795-006A-126  | Sequence 126, Appl   |
| 42 | 40   | 3.1 | 32204 | 10 | US-09-764-846-306   | Sequence 306, Appl   |
| 43 | 39.6 | 3.1 | 460   | 10 | US-09-960-352-10251 | Sequence 10251, Appl |
| 44 | 39.6 | 3.1 | 6503  | 10 | US-09-935-541-12    | Sequence 12, Appl    |
| 45 | 39.4 | 3.1 | 2910  | 10 | US-09-833-790-418   | Sequence 418, Appl   |

```

RESULT 1
US-09-844-864-1
; Sequence 1, Application US/09844864
; Patent No. US20020042926A1
GENERAL INFORMATION:
APPLICANT: Matczuk, Martin
APPLICANT: Ren, Yongsheng
APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / CTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIORITY FILING DATE: 1998-10-20
PRIORITY FILING DATE: 1998-10-28
PRIORITY APPLICATION NUMBER: PCT/US99/25209
PRIORITY FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-844-864-1

Query Match          100.0%; Score 1277; DB 10; Length 1277;
Best Local Similarity 100.0%; Pred. No. 2e-307;
Matches 1277; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY      1  AAGCGGGGCGAGGCGGGGAGCGACCCATGTCCTCCGGCGAGCAGCTTCCACCCCTGCGCG 60
Db      1  AAGCGGGGCGAGGCGGGGAGCGACCCATGTCCTCCGGCGAGCAGCTTCCACCCCTGCGCG 60

QY      61  CATCTTATTCGAGCGACCAAAAGCGGGGATGCTGTGAGGTTGCGAGCCAGGGCGTC 120
Db      61  CATCTTATTCGAGCGACCAAAAGCGGGGATGCTGTGAGGTTGCGAGCCAGGGCGTC 120

QY      121  CGAAGCGGCGCCCTCTCTCTCCCGCGCTACAGACGCTCATGCGCGGGGAGTACGTC 180
Db      121  CGAAGCGGCGCCCTCTCTCTCCCGCGCTACAGACGCTCATGCGCGGGGAGTACGTC 180

QY      181  GACAGCCACGAGGCGGACAGCTCATGCGCTGTGTCGCGATGGGTGCCCGGTGCTC 240
Db      181  GACAGCCACGAGGCGGACAGCTCATGCGCTGTGTCGCGATGGGTGCCCGGTGCTC 240

```

181 GACAGCCACAGGAGGACACAGCTCATGGCCCTGCTGTGCGAGATGGATCCCGGTGCTC 240  
 QY 241 AGCAGCCGTGAGGCTGCGGTGCAAGTGAACCCGGCGCCGAGCCTCGGTGCAAGTGTCA 300  
 Db 241 AGCAGCCGTGAGGCTGCGGTGCAAGTGAACCCGGCGCCGAGCCTCGGTGCAAGTGTCA 300  
 QY 301 CTGCGGCGCGGACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 360  
 Db 301 CTGCGGCGCGGACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 360  
 QY 361 TCCTGTCAACCCCGTGGCCAGCGCGCGCGGAGATCCCGCGATCCTGCGAGACCGTA 420  
 Db 361 TCCTGTCAACCCCGTGGCCAGCGCGCGCGGAGATCCCGCGATCCTGCGAGACCGTA 420  
 QY 421 GCGCCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 Db 421 GCGCCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 481 AGGAGACACCCCGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 Db 481 AGGAGACACCCCGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 QY 541 CCGAGAGAGGTGGCCCGAG 600  
 Db 541 CCGAGAGAGGTGGCCCGAG 600  
 QY 601 CAGGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
 Db 601 CAGGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
 QY 661 GCGGAGTGTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
 Db 661 GCGGAGTGTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
 QY 721 GACCCGCGTGTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
 Db 721 GACCCGCGTGTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
 QY 781 AAGGAGGCGCTGCGGTTCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 Db 781 AAGGAGGCGCTGCGGTTCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 QY 841 TGCAGAGTGTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900  
 Db 841 TGCAGAGTGTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900  
 QY 901 CTTCAGAGTGTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
 Db 901 CTTCAGAGTGTGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
 QY 961 CACCTGTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020  
 Db 961 CACCTGTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020  
 QY 1021 CCGTAAAGCGCCGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080  
 Db 1021 CCGTAAAGCGCCGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080  
 QY 1081 CGAGAGACCTTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGT 1140  
 Db 1081 CGAGAGACCTTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGT 1140  
 QY 1141 TGGGCGTAAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200  
 Db 1141 TGGGCGTAAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200  
 QY 1201 TCTTCATGAGAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 Db 1201 TCTTCATGAGAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 QY 1261 AAAAAA 1277  
 Db 1261 AAAAAA 1277

RESULT 2  
 US-09-844-864-18  
 ; Sequence 18, Application US/09844864  
 ; Patent No. US20020042926A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Matzuk, Martin  
 ; APPLICANT: Ren, Yongsheng  
 ; APPLICANT: Wu, Xuemei  
 ; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS  
 ; FILE REFERENCE: P01925US2 / 09807797 / OFA 99-48  
 ; CURRENT APPLICATION NUMBER: US/09/844,864  
 ; CURRENT FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: 60/106,020  
 ; PRIOR FILING DATE: 1998-10-28  
 ; PRIOR APPLICATION NUMBER: PCT/US99/25209  
 ; PRIOR FILING DATE: 1999-10-28  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 18  
 ; LENGTH: 814  
 ; TYPE: DNA  
 ; ORGANISM: mus musculus  
 US-09-844-864-18

Query Match 62.6% Score 799.4; DB 10; Length 814;  
 Best Local Similarity 99.9%; Pred. No. 4,7e-189;  
 Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGGCGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 62  
 Db 1 GCGGCGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60  
 QY 63 TCCTTATCCGAGGCGCACCAAGCCGGGAGTGGCTGGAAGTTTGGAGGCGAGGGGCTCCG 122  
 Db 61 TCCTTATCCGAGGCGCACCAAGCCGGGAGTGGCTGGAAGTTTGGAGGCGAGGGGCTCCG 120  
 QY 123 ACCGCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 182  
 Db 121 ACCGCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 183 CAGCCAGCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 242  
 Db 181 CAGCCAGCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
 QY 243 CAGCCGTAAGCTGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 302  
 Db 241 CAGCCGTAAGCTGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
 QY 303 CCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 362  
 Db 301 CCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
 QY 363 CTGTAAACCCGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 422  
 Db 361 CTGTAAACCCGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
 QY 423 CCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482  
 Db 421 CCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 483 GCGAGACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 542  
 Db 481 GCGAGACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 QY 543 GAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602  
 Db 541 GAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 QY 603 GCGTCAAGGCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 662  
 Db 601 GCGTCAAGGCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660

*Accepted*



```

: RESULT 6
: US-09-844-864-19
: Sequence 19, Application US/09844864
: Patent No. US20020042926A1
: GENERAL INFORMATION:
: APPLICANT: Matzuk, Martin
: APPLICANT: Ren, Yongsheng
: APPLICANT: Wu, Xuemei
: TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
: FILE REFERENCE: P01925052 / 09807797 / OTA 99-48
: CURRENT APPLICATION NUMBER: US/09/844,864
: CURRENT FILING DATE: 2001-04-27
: PRIOR APPLICATION NUMBER: 60/106,020
: PRIOR FILING DATE: 1998-10-28
: PRIOR APPLICATION NUMBER: PCT/US99/25209
: PRIOR FILING DATE: 1999-10-28
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 19
: LENGTH: 123
: TYPE: DNA
: ORGANISM: Mus musculus

```

```

US-09-844-864-20
? Sequence 20, Application US/09844864
? Patent No. US20020042926A1
? GENERAL INFORMATION:
? APPLICANT: Matczuk, Martin
? APPLICANT: Ren, Jiongsheng
? APPLICANT: Wu, Xuemei
? TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
? FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
? CURRENT APPLICATION NUMBER: US/09/844,864
? PRIOR FILING DATE: 2001-04-27
? PRIOR APPLICATION NUMBER: 60/106,020
? PRIOR FILING DATE: 1998-10-28
? PRIOR APPLICATION NUMBER: PCT/US99/25209
? PRIOR FILING DATE: 1999-10-28
? NUMBER OF SEQ ID NOS: 25
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 20
? LENGTH: 105
? TYPE: DNA

```



```

: RESULT 10
: US-09-954-456-1601
: Sequence 1601. Application US/09954456
: Patent No. US20020115057A1
: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
: TITLE OF INVENTION: Set
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637

```

|                           |        |                  |           |              |
|---------------------------|--------|------------------|-----------|--------------|
| Query Match               | 3.88;  | Score 49;        | DB 12;    | Length 7177; |
| Best Local Similarity     | 48.08; | Pred. No. 0.016; |           |              |
| Matches 201; Conservative | 0;     | Mismatches 215;  | Indels 3; | Gaps 2       |

|    |      |   |      |
|----|------|---|------|
| QY | 250  | GACCTCTCGSGTGCAGGTGMAACCCCGCGCGCGAGCGCGCTCGTGTCACTGTTTCACTCGAGGCG | 309  |
| Db | 2753 | GAGGCCCCGAGAGCGGAGAGACGGGGAGGCCGGGTGCCCGGGAGAGACGGCCGGCGCCGAC     | 2812 |
| QY | 310  | CGCAGCTGTGAGCGCTGCAGAGGTGTCCTCAGCGCACGCCCGGACCCGATCGAGTTCTGTCA    | 369  |
| Db | 2813 | CGCA - GCCACAGCAAGSAGSGCCGCGGGGCC - - CCGGAGAGCGCGGAGAGCGCGCGGA   | 2869 |
| QY | 370  | CCCCGTGGCCACGCGCGCGCGCGGAGATCCCGCATCTGTGGCAGACCGTACCCCGTTC        | 429  |
| Db | 2870 | GGCCCAAGCCCCCGAGGCGCGCGCGCGCACACACGGCGCGGCTCCCGCGGAGAGCGCGCC      | 2929 |
| QY | 430  | TCGTCTCGTGACCTTTCTGTGCGCTCTCTCTCTCTACTGTGCGGAGTTCAGCAGCAGACA      | 489  |
| Db | 2930 | GAGCGGGAGGCCCGACGCCACCGCGGCACCGGCAACGAGATCCAGCAAGSAGTGGCCG        | 2989 |
| QY | 490  | CCCAACGAAGAGAGAGGGGAGACCCCGCATCTCTCGGAGACCCGGGAACCGGACCGAGAGAG    | 549  |
| Db | 2990 | GGCGCCAAAGGCGAGCGCGCGCGCGGACACCGCGGCGCCCGGACGGGCGCCCGGAG          | 3049 |
| QY | 550  | GTGGCCCGCAGAGAAAGCGGTCTCCCCAGCCCGCGAAGCGSAGSAGGCGATTTAGCGTGA      | 609  |
| Db | 3050 | GGCGGAGACCGGGGAGGAGCCGGCGCGCGACCCGGGACCCGCGACAAAGCGCAGCGCTGT      | 3109 |
| QY | 610  | GGGCGAGGCGGGGTGGGAGCAGCAACCCACCGCGAGAGCCGGAACAGTGTGGCGCGCAT       | 668  |
| Db | 3110 | CACGAGGCTGTGGAGAAAGAGACCAACGAGGAAGSAGGCCACGAGAGAGSAGGCTGAGAT      | 3168 |

```

RESULT 12
US-09-954-456-1179
; Sequence 1179, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954, 456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233, 617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234, 052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234, 923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235, 134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235, 637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235, 638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235, 711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 1179
; LENGTH: 7364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1179

Query Match 3 8%; Score 49; DB 10; Length 7364;
Best Local Similarity 48.08; Pred. No. 0.016;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

```

|    |      |  |      |
|----|------|--|------|
| OY | 250  | GACCTGCGGTGAGTGTACACCCGCGCGGACGCCCTCTGCTGCAATGTTTCACGTCCGGGGCC | 309  |
|    |      |  |      |
| Db | 2753 | GAGACCCCGGAAAGCGAGAGCGGGGAGCCCGGTGTCCCGGAGAGGCGCGCGCCGAC       | 2812 |
|    |      |  |      |
| OY | 310  | CGCAGCGTGCAGCTTCACGAGTGTCCGAGCACAGCCCGACGCCGATCGGTGTTCTGTCA    | 369  |
|    |      |  |      |
| Db | 2813 | CGCA -GCCACAGCAAGAGAGCCGCGGGGCG -CCCGAGAGCGCGAGCGCGCGCA        | 2869 |
|    |      |  |      |
| OY | 370  | CCCCGTGGCCAGCGCGCGCGCGGAGATCCCGCGATCTGTGCAACCGTAGCCCGCTTC      | 429  |
|    |      |  |      |
| Db | 2870 | GGCCCAAGGCCCGGAGGCGCGCGCGGACCAACCGCGCGGCTTCGCCGGAGAGCGGCC      | 2929 |
|    |      |  |      |
| OY | 430  | TGCTGCTGTACCTTCTGTGGCCCTCTCTCTACCTGTGAGGTTCGGGAGGACAGGACAGACA  | 489  |
|    |      |  |      |
| Db | 2930 | GAGCGGAGCGCCCGAGCGCACCGCGCGCACCGGACACAGATCCGACACAGAGTGCGCC     | 2989 |
|    |      |  |      |
| OY | 490  | CCCAACGAGGAGAGAGGGGAGACCCGGCATCTTCGGGGACCCGGGAACTCGAGCCGAAGAG  | 549  |
|    |      |  |      |
| Db | 2990 | GGCGCCAAAGGGCGAGCGCGCGCGCGCGCACCGCGCGCGCCCGCCAGCGGGGGCCCGGAG   | 3049 |
|    |      |  |      |
| OY | 550  | GTCGGCCCGCAGAGAAACGCGTCCCGACGCGCGGAACCGAGGAGGGCGATGTTTCAGCGTGA | 609  |
|    |      |  |      |
| Db | 3050 | GCGGAGAGCGGGGGAGGAGCCGGCGGGCGGACCGCGGCGCCGACAGGGCGGACCTGCT     | 3109 |
|    |      |  |      |
| OY | 610  | GGGCAACCGGGGTGTGGAGACGACAGCCACCCGAGAGGACCGGAACTGTGGCGCGCAT     | 668  |
|    |      |  |      |
| Db | 3110 | CACGAGGCTGTGGAGAAAGAGACACAGGAGAGGAGGCCACGAGGAAGGAGGCTGTGAGAT   | 3168 |

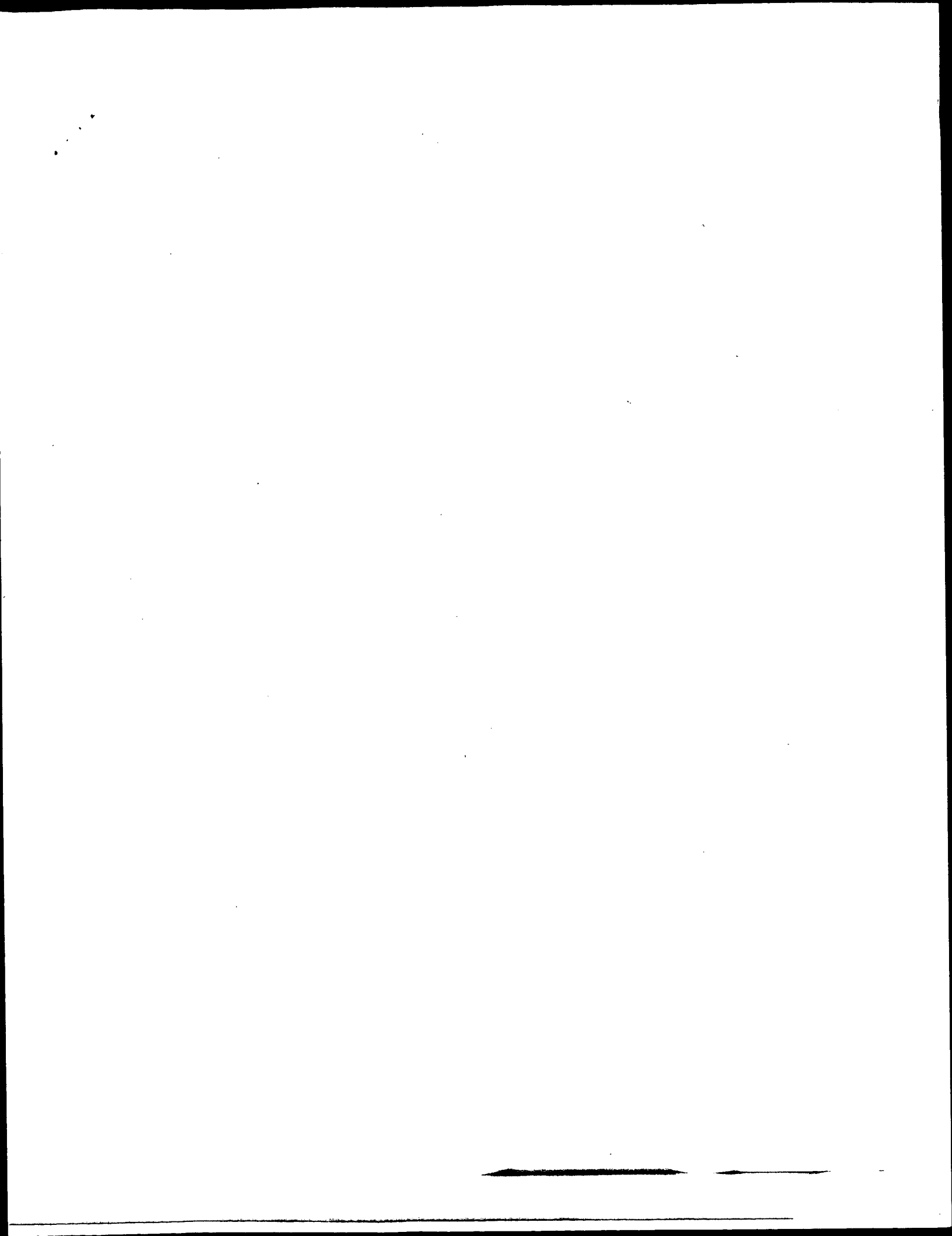
```

RESULT 13
US-10-033-026-5
/ Sequence 5, Application US/10033026
/ Patent No. US20020147309A1
/ GENERAL INFORMATION:
/ APPLICANT: Lipscombe, Diane
/ APPLICANT: Schorge, Stephanie
/ TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
/ FILE REFERENCE: B1055/7000
/ CURRENT APPLICATION NUMBER: US/10/033,026
/ CURRENT FILING DATE: 2001-12-28
/ PRIOR APPLICATION NUMBER: US 09/268,163
/ PRIOR FILING DATE: 1999-03-12
/ PRIOR APPLICATION NUMBER: US 60/077,901
/ PRIOR FILING DATE: 1998-03-13
/ NUMBER OF SEQ ID NOS: 28
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5
/ LENGTH: 7364
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 146..7162
/
US-10-033-026-5

```

[illegible]





OM nucleic - nucleic search, using sw model

Run on: January 18, 2003, 09:43:41 ; Search time 296 Seconds

(without alignments)  
9715.547 Million cell updates/sec

Title: US-09-844-864-1

Perfect score: 1277

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
post-processing: Minimum Match 0%
```

Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

|  |     |   |
|--|-----|---|
|  | 1:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT*  |
|  | 2:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT*  |
|  | 3:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT*  |
|  | 4:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT*  |
|  | 5:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT*  |
|  | 6:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT*  |
|  | 7:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT*  |
|  | 8:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT*  |
|  | 9:  | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT*  |
|  | 10: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT*  |
|  | 11: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT*  |
|  | 12: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT*  |
|  | 13: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT*  |
|  | 14: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT*  |
|  | 15: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT*  |
|  | 16: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT*  |
|  | 17: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT*  |
|  | 18: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT*  |
|  | 19: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT*  |
|  | 20: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT*  |
|  | 21: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT*  |
|  | 22: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT* |
|  | 23: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT* |
|  | 24: | /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT*  |

## SUMMARIES

| Result<br>No. | Score | Query | Match  | Length | DB       | ID | Description           |
|---------------|-------|-------|--------|--------|----------|----|-----------------------|
| 1             | 1265  | 99.1  | 1276   | 21     | AA000234 |    | Muscle cdc2cye-3 gene |
| 2             | 56.8  | 4.6   | 14955  | 20     | AA053491 |    | Human adenovine A1    |
| 3             | 52    | 4.1   | 114955 | 20     | AA053491 |    | Human adenovine A1    |
| 4             | 49.6  | 3.9   | 400    | 19     | AA064545 |    | M. tuberculosis in    |
| 5             | 49.6  | 3.9   | 400    | 19     | AA044436 |    | Mycobacterium tuber   |
| 6             | 49.6  | 3.9   | 400    | 20     | AA021934 |    | M. tuberculosis an    |
| 7             | 49.6  | 3.9   | 400    | 20     | AA021934 |    | M. tuberculosis re    |
| 8             | 49.6  | 3.9   | 14800  | 24     | AA066291 |    | Lung cancer relate    |
| 9             | 49.6  | 3.9   | 14835  | 24     | AA059485 |    | Human DNA sequenc     |

|   |    |      |     |       |    |           |    |      |     |       |    |          |
|---|----|------|-----|-------|----|-----------|----|------|-----|-------|----|----------|
| C | 26 | 46.6 | 3.6 | 7.45  | 15 | AA0C28237 | 44 | 45.8 | 3.6 | 6.232 | 13 | AA029289 |
| C | 27 | 46.6 | 3.6 | 7.45  | 17 | AA0F80777 | 45 | 45.8 | 3.6 | 6.232 | 13 | AA029289 |
| C | 28 | 46.6 | 3.6 | 7.45  | 17 | AA0F80777 | 45 | 45.8 | 3.6 | 6.232 | 13 | AA029289 |
| C | 29 | 46.6 | 3.6 | 7.45  | 17 | AA0F80777 | 45 | 45.8 | 3.6 | 6.232 | 13 | AA029289 |
| C | 30 | 46.6 | 3.6 | 7.45  | 17 | AA0F80777 | 45 | 45.8 | 3.6 | 6.232 | 13 | AA029289 |
| C | 31 | 46.6 | 3.6 | 7.45  | 20 | AA022060  | 46 | 46.6 | 3.6 | 6.232 | 13 | AA029289 |
| C | 32 | 46.6 | 3.6 | 7.45  | 20 | AA022060  | 46 | 46.6 | 3.6 | 6.232 | 13 | AA029289 |
| C | 33 | 46.6 | 3.6 | 7.45  | 22 | AA0X01175 | 47 | 47.6 | 3.6 | 6.232 | 13 | AA029289 |
| C | 34 | 46.6 | 3.6 | 7.45  | 22 | AA0X01175 | 47 | 47.6 | 3.6 | 6.232 | 13 | AA029289 |
| C | 35 | 46.6 | 3.6 | 10.06 | 22 | AA0F53050 | 48 | 48.2 | 3.6 | 6.232 | 13 | AA029289 |
| C | 36 | 46.6 | 3.6 | 10.06 | 22 | AA0F53050 | 48 | 48.2 | 3.6 | 6.232 | 13 | AA029289 |
| C | 37 | 46.6 | 3.6 | 19.66 | 20 | AA022061  | 49 | 49   | 3.6 | 6.232 | 13 | AA029289 |
| C | 38 | 46.6 | 3.6 | 19.66 | 22 | AA0C0318  | 50 | 50   | 3.6 | 6.232 | 13 | AA029289 |
| C | 39 | 46.6 | 3.6 | 20.03 | 15 | AA062838  | 51 | 51   | 3.6 | 6.232 | 13 | AA029289 |
| C | 40 | 46.6 | 3.6 | 20.03 | 17 | AA058339  | 52 | 52   | 3.6 | 6.232 | 13 | AA029289 |
| C | 41 | 46.6 | 3.6 | 20.03 | 17 | AA0F80900 | 53 | 53   | 3.6 | 6.232 | 13 | AA029289 |
| C | 42 | 46.6 | 3.6 | 20.03 | 17 | AA0F80900 | 53 | 53   | 3.6 | 6.232 | 13 | AA029289 |
| C | 43 | 46.6 | 3.6 | 20.03 | 20 | AA06739   | 54 | 54   | 3.6 | 6.232 | 13 | AA029289 |
| C | 44 | 46.6 | 3.6 | 20.03 | 20 | AA052062  | 55 | 55   | 3.6 | 6.232 | 13 | AA029289 |
| C | 45 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 46 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 47 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 48 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 49 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 50 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 51 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 52 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 53 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 54 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 55 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 56 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 57 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 58 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 59 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 60 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
| C | 61 | 46.6 | 3.6 | 20.03 | 22 | AA0C9031  | 56 | 56   | 3.6 | 6.232 | 13 | AA029289 |
|   |    |      |     |       |    |           |    |      |     |       |    |          |

## ALIGNMENTS

|  |
|--|
| Human neuronal calcium-binding protein 1 |
| DNA encoding human                       |
| Human calcium channel gamma-1            |
| Sequence encoding                        |
| Human neuronal calcium-binding protein 1 |
| DNA encoding human                       |
| Human calcium channel gamma-1            |
| Lung cancer related                      |
| N-type calcium channel alpha-1A          |
| Micromonospora DNA                       |
| Human CDNA 5'-end                        |
| Human CDNA clone r                       |
| Human full-length                        |
| Chemically treated                       |
| Human immune system                      |
| Gcf gene coding seq                      |
| GCF gene coding seq                      |
| Human glial growth factor                |
| Human glial growth factor                |
| Human glial growth factor                |
| Human MTH2 relate                        |
| Human neuroglial G                       |
| I36448 cDNA clone.                       |
| Human NR1 exon si                        |
| Human NR1 exon si                        |
| Human MTH2 relate                        |
| L1260 cDNA clone.                        |
| GGF2IBBS. Homo sa                        |
| GGF2IBBS. Homo sa                        |
| Human neuroglial G                       |
| Glial growth facto                       |
| Glial growth facto                       |
| Human MTH2 relate                        |
| I36532 cDNA clone.                       |
| Human calcium chan                       |

|          |  |
|----------|--|
|          | RESULT 1   |
| AAD00294 | ID AAD00294 standard; cDNA: 1276 BP.                                     |
| XX AC    | AAD00294;  |
| XX AC    | 05-SEP-2000 (first entry)  |
| XX DT    | Mouse oocyte-specific OI-180 cDNA clone.                                 |
| XX DE    | Oocyte-specific: ovary; OI-180; mouse; gynaecological; treatment; screen |
| XX XX    | cell proliferative disorder; cell degenerative disorder; contraceptive;  |
| KW KW    | modulator; signalling pathway; human infertility; cancer; ovulation; ss. |
| XX OS    | Mus sp.  |
| XX FH    | Key Location/Qualifiers  |
| FH FT    | 28..1113   |
| FT CDS   | /tag= a  |
| FT FT    | /product= "Mouse oocyte-specific protein, OI-180"                        |
| XX NM    | WO200024735-A1.  |
| NM PD    | 04-MAY-2000.   |
| XX XX    |  |
| XX PF    | 28-OCT-1999; 99WO-US25209.   |
| XX PR    | 28-OCT-1998; 98US-0106020.   |
| XX PA    | (BAYU ) BAYLOR COLLEGE MEDICINE.   |
| XX PI    | Matzuk KM, Wang P;   |





oligonucleotides (specifically AAX5180-271) can be used for the antisense treatment of diseases and conditions. Typical diseases and conditions are those associated with impaired respiration and inflammation, including lung diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer, pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as well as all types of cancers which may metastasize or have metastasized to the lungs, including breast and prostate cancer.

Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other:

Query Match 4.1%; Score 52; DB 20; Length 114955;  
Best Local Similarity 32.3%; Pred. No. 0.054;  
Matches 244; Conservative 50; Mismatches 456; Indels 5; Gaps 2;

QY 13 GCGGCGAGACGACCCATGTTCCCGGAGACAGTTCACCCCTGCCGATCTATCGG 72  
DB 111059 GCGCGGAGAGCCGNNHNNNSGCCGCAAGCCAGCGGCGNNHNNNSGCCGACGCG 111000  
QY 73 CAGGCGACCAAGCCGGGATGCTGAGAGTTCGAGCCAGGAGGCTCCGACCCGCGCC 132  
DB 110999 GCGCGAGCCGNNHNNNSGCCGCGCAAGAGGSCCGGACCGCGGCGGCGNNHNNNSGCC 110940  
QY 133 CCGTCCTTCTCTCC---GCTACAGACAGCTCATGCGCCGAGTACGTGACAGCCA 188  
DB 110939 CCGCGCGGCGCCGCGCGGCAANNHNNNSTCCGCGAGCGCGCGCGAGCACCCNNHNN 110880  
QY 189 CCAGCGGAGACAGCTCATGCGCTGTCGCGGATGGGTCCCGGATGAGCAGCGC 248  
DB 110879 NSGGCGCGCGAGGCGCCACCCCGCCNNHNNNSGCCCGCGAGGAGGCGCGGCG 110820  
QY 249 TGACGCTGCGGTGACAGTGAACCCGCG-CCGCGAGCGCTCGGTGAGTTCATCTCGGCG 307  
DB 110819 CNHNNNSGCGCGCGCGCGCGCGGCGGCGNNHNNNSGCGAGCGCGCGAGCGGCG 110760  
QY 308 GCGCGACGCTGACAGCTGAGGAGTCCGAGCCGCGCGCGCGGATCGGTTCTCTGTC 367  
DB 110759 GAGCGGAGCGCGCGNNHNNNSGCGCGCGCGCGCGCGCGCGCGCGNNHNNNSGCG 110700  
QY 368 AACCCGCTGAGCGCGCGCGCGCGGAGATCCCGGATCTGAGAGACCTGACCCCGCT 427  
DB 110699 GCG 110640  
QY 428 TCTGCTGCTGAGCTTCTGTGCGCTCTCTCTCACTGAGAGTTCGCGGAGCGAGCGA 487  
DB 110639 GCGCGCGCGCGNNHNNNSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110580  
QY 488 CACCCAGAGAGGAGAGGAGCGCGCGCATCTCGGCGGAGCGCGGAGAACCGCGAGAG 547  
DB 110579 GGGCGGCGCGGNNHNNNSGCGCGCGCGCGCGGNNHNNNSGCGCGCGCGCGCG 110520  
QY 548 AGGTGCGCGGAGAGAAAGCGGTCCCGAGCGGAGAGAGAGGCGCATGTTCAAGCTG 607  
DB 110519 NHHNNNSGCG 110460  
QY 608 CAGGCGAGCGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 667  
DB 110459 CGGCGCGCGCGNNHNNNSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110400  
QY 668 TGTAGTGTGAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 727  
DB 110399 GCGCGGCG 110340  
QY 728 GTGATTGAGTGCCTCTGAGAGACAGCGCTCCCGC 762  
DB 110339 NHHNNNSGCG 110305

RESULT 4

AAV64545  
ID AAV64545 standard; DNA; 400 BP.

XX AAV64545;

XX 27-JAN-1999 (first entry)

DE M. tuberculosis immunogenic polypeptide XP22 5'-end DNA.

XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;  
KW vaccine; pharmaceutical; infection; diagnosis; ss.

OS Mycobacterium tuberculosis.

PN W09B16646-A2.

PD 23-APR-1998.

PF 07-OCT-1997; 97WO-US18293.

PR 13-MAR-1997; 97US-0818112.

PR 11-OCT-1996; 96US-0730510.

XX (CORI-) CORIXA CORP.

PI Campos-Neto A, Dillon DC, Houghton R, Iodes MJ;  
PI Reed SG, Skeiky YAM, Twardzik DR, Vedvick TS;

DR WPI; 1998-261042/23.

PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used  
PT to develop products for the detection of M. tuberculosis infection  
PT and for diagnosis, treatment and prevention of tuberculosis

XX Claim 4; Page 175; 230pp; English.

XX This sequence encodes an immunogenic portion of a soluble Mycobacterium  
XX tuberculosis (MT) antigen which can be used in a method for inducing  
XX protective immunity against tuberculosis (TB). This sequence can be  
XX formulated into vaccines and/or pharmaceutical compositions for  
XX immunizing against M. tuberculosis infection or may be used for the  
XX diagnosis of tuberculosis.

XX Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;

Query Match 3.9%; Score 49.6; DB 19; Length 400;  
Best Local Similarity 49.2%; Pred. No. 0.033;  
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 458 CTTCACTGAGAGTTCGCGGAGGCGAGACACACCCAGAGAGGAGAGGAGCGCCGAT 517  
DB 119 CTTCACTGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAA 178  
QY 518 CTTGCGGAGCGCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 577  
DB 179 CCGCGGAGAGCGCGGAGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 238  
QY 578 CCGGAGAGGAGAGGCGGATGTTCAAGCTGCAAGGCGAGCGCGCGGTGGAGAGCAGC 637  
DB 239 CCGGAGAGGCG 298  
QY 638 CACCGAGAGCGCGGAGACAGTGTGGCGCGGATGCAAGTCTGAGCCGCGGAGAGAG 697  
DB 299 CCGGAGAGCG 358  
QY 698 GTCTGCGCGAGAGATGGCTAGG 721  
DB 359 CCGCGGAGAGGAGTGCATCTCAGG 382

RESULT 5  
AAV44436  
ID AAV44436 standard; DNA; 400 BP.



XX AAV4436;  
AC  
XX  
DT 09-NOV-1998 (first entry)  
XX  
DE Mycobacterium tuberculosis antigen XP22 5' DNA.  
XX  
XX Tuberculosis; infection; diagnosis; antigen: XP22; ss.  
XX  
OS Mycobacterium tuberculosis strain Erdman.  
XX  
PN W09816645-A2.  
XX  
PD 23-APR-1998.  
XX  
PF 07-OCT-1997; 97WO-US18214.  
XX  
PR 13-MAR-1997; 97US-0818111.  
PR 11-OCT-1996; 96US-0729622.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;  
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;  
XX  
XX WPI: 1998-251292/22.  
XX  
XX New isolated Mycobacterium tuberculosis polypeptides and DNA - used  
PT to develop products for the detection of M. tuberculosis infection  
PT and diagnosis of tuberculosis  
XX  
XX Claim 4: Page 185-186; 250pp; English.  
XX  
XX This is the 5' region of DNA coding for an antigenic portion of  
CC Mycobacterium tuberculosis antigen XP22. 3' DNA is provided in  
CC AAV4437. XP22 DNA was isolated from a M. tuberculosis strain Erdman  
CC genomic DNA expression library using sera from patients having  
CC extrapulmonary tuberculosis. It bears no similarity to known  
CC sequences. The invention relates to methods and compositions for  
CC diagnosing tuberculosis. It provides polypeptides (see  
CC AAV64391-W64379) comprising an antigenic portion of a soluble M.  
CC tuberculosis antigen, or an immunogenic portion of a M.  
CC tuberculosis antigen, as well as DNA sequences encoding such  
CC polypeptides, recombinant expression vectors and transformed or  
CC transduced host cells. Also claimed are methods and diagnostic  
CC kits for detecting M. tuberculosis infection in a patient using  
CC these polypeptides, antibodies or oligonucleotide probes and  
CC primers.  
XX  
XX Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;  
SQ

Query Match 3.9%; Score 49.6; DB 19; Length 400;  
Best Local Similarity 49.2%; Pred. No. 0.033;  
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 458 CCTACTGAGAGTTGCCGGAGCGACGACACCCAGAGAGAGAGAGCCCGGCAT 517  
DB 119 CTTTACCCAAAGCGCGCCGACGCAACGCGGCAACGCGGTGACGGCGGGCGCAA 178

QY 518 CCTCGGGAGCCCGGAGCGAGCGAGAGTGGCCGGAGGAAAGCGGTCCCGCAGC 577  
DB 179 CGGCGGAAAGCGCGGAAACCGCGACAGACACCAACCGCGCGCGCCGACACAGG 238

QY 578 CGCGAAGCGAGAGCGGAGTGTTCAGGCTGACAGGCGAGCGGGGTGGAGACAGCAC 637  
DB 239 CGGCGAGCGGGGCGCGGGGCGCGGACCGGAGAAAGCGGGAGACCGCGCGAC 238

QY 638 CACCGAGAGCGGAAAGTGTGCGCGCGGAGTGCAGTGTGAGCTGGAGCGAGAGCCAT 697  
DB 299 CGGACCGCGCGCGGCGGAAAGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGAC 358

QY 698 GTCTGCCGAGAGATGGCTCAGG 721  
II III

DB 359 CGGCGGCGAGCGTGCCTCTCAGG 382

RESULT 6  
AA219346  
ID AA219346 standard; DNA: 400 BP.  
XX  
XX AA219346;  
XX  
XX 05-NOV-1999 (first entry)  
XX  
XX M. tuberculosis antigen 5' XP22 DNA sequence.  
XX  
XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;  
KW immunotherapy; diagnosis; immunisation; vaccine; infection;  
KW immune response; skin test; ss.  
XX  
XX Mycobacterium tuberculosis.  
XX  
XX W09942076-A2.  
XX  
XX 26-AUG-1999.  
XX  
XX 17-FEB-1999; 99WO-US03268.  
XX  
XX 05-MAY-1998; 98US-0072967.  
XX 18-FEB-1998; 98US-0025197.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;  
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;  
XX  
XX WPI: 1999-527409/44.  
XX  
XX New antigens from Mycobacterium tuberculosis useful in diagnostic  
PT skin tests and protective or therapeutic vaccines or compositions  
PT  
XX  
XX Claim 4: Page 170; 299pp; English.  
XX  
XX The present invention describes polypeptides comprising an immunogenic  
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described  
CC are vaccines and fusion protein containing M. tuberculosis Ag's.  
CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and  
CC other polypeptides fragments, can be used in pharmaceutical compositions  
CC or vaccines to generate a protective or therapeutic immune response to  
CC M. tuberculosis and as reagents in skin tests for diagnosis of  
CC tuberculosis. Ag can induce proliferation of, or cytokine secretion  
CC by, T, B or natural killer cells and/or macrophages in  
CC tuberculosis-immune subjects. AA219249 to AA219460 and AAV39083 to  
CC AA39225 are used in the exemplification of the present invention.  
XX  
XX Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;  
SQ

Query Match 3.9%; Score 49.6; DB 20; Length 400;  
Best Local Similarity 49.2%; Pred. No. 0.033;  
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 458 CCTACTGAGAGTTGCCGGAGCGACACACCCAGAGAGAGAGAGCCCGGCAT 517  
DB 119 CTTTACCCAAAGCGCGCGCGACGCAACGCGGCAACGCGGTGACGGCGGGCGCAA 178

QY 518 CCTCGGGAGCCCGGAGCGGAGCGAGAGTGGCCGGAGGAAAGCGGTCCCGCAGC 577  
DB 179 CGGCGGAAAGCGCGGAAAGCGCGACAGACCAACCAACCGCGCGCGCGACAGG 238

QY 578 CGCGAAGCGAGAGCGGAGTGTTCAGGCTGACAGGCGAGCGGGGTGGAGCGAGCCAT 637  
DB 239 CGGCGAGCGGGGCGCGGGGCGCGGAGAACCGCGGAGAACCGCGCGCGGCGAC 298

QY 638 CACCGAGAGCGGAAAGTGTGCGCGGAGTGCAGTGTGAGCTGGAGCGAGAGCCAT 697  
DB 299 CGGACCGCGCGCGGCGGAAAGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGAC 358



XX WPI: 2002-188264/24.  
 XX  
 XX Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set -  
 XX  
 PS Claim 1: SEQ ID 4628; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in AB161664  
 CC to AB170110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilms' tumour.  
 CC  
 SO Sequence 14800 BP; 2890 A; 4567 C; 5319 G; 2024 T; 0 other;  
 Query Match 3.9%; Score 49.6; DB 24; Length 14800;  
 Best Local Similarity 48.3%; Pred. No. 0.1;  
 Matches 139; Conservative 0; Mismatches 149; Indels 0; Gaps 0;  
 Oy 507 GAGCCCGCATCTCTGGGAGCCCGGAGACCGGAGAGAGTGGCCGGAGGAAAGC 566  
 Db 4122 GAGCCAGTACATCAATGATTCATGACGACATCTGCGGCCATGAGAGAGAGAGCT 4181  
 Oy 567 GGTCCCGCAGCCGAGAGGAGAGGAGGATTTGAGCTGAGGAGGAGGAGGAGGAG 626  
 Db 4182 GGTCTAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4241  
 Oy 627 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 686  
 Db 4242 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4301  
 Oy 687 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 746  
 Db 4302 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4361  
 Oy 747 AGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 794  
 Db 4362 CGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4409  
 RESULT 9  
 AAS94858  
 ID AAS94858 standard; DNA; 14835 BP.  
 XX  
 AC AAS94858;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human DNA sequence #113 expressed during foam cell differentiation.  
 XX  
 KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
 KW cardiovascular disorder; coronary artery disease; gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 MO200177389-A2.  
 XX  
 PN 18-OCT-2001.  
 PD  
 XX

PF 04-APR-2001; 2001WO-US111128.  
 XX  
 XX 05-APR-2000; 2000US-195106P.  
 PR  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 XX Shiftman D, Somogyi R, Lawn R, Sellhammer JJ, Porter GJ, Mikita T;  
 PI Tal J;  
 DR WPI: 2002-010925/01.  
 XX  
 DR Composition useful for diagnosis of conditions, disorders or diseases  
 PT associated with atherosclerosis, comprises several polynucleotides that  
 PT are differentially expressed in foam cell development -  
 XX  
 PS Claim 1: Page 165-169; 315pp; English.  
 XX  
 CC The present invention relates to the isolation of human polynucleotide  
 CC sequences that are differentially expressed during foam cell  
 CC differentiation. The polynucleotide sequences of the invention or a  
 CC composition comprising these polynucleotides are useful as a high  
 CC throughput method for detecting altered expression of one or more  
 CC polynucleotides in a sample. The polynucleotides can be used in the  
 CC diagnosis of disorders associated with foam cell development such as  
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
 CC coronary artery disease. The polynucleotide sequences can also be used  
 CC as PCR primers and probes. The polynucleotides of the invention are also  
 CC useful in gene therapy. AAS94746-AAS95021 represent the human  
 CC polynucleotide sequences of the invention which are differentially  
 CC expressed during foam cell differentiation.  
 CC  
 SO Sequence 14835 BP; 2881 A; 4568 C; 5306 G; 2018 T; 62 other;  
 Query Match 3.9%; Score 49.6; DB 24; Length 14835;  
 Best Local Similarity 48.3%; Pred. No. 0.1;  
 Matches 139; Conservative 0; Mismatches 149; Indels 0; Gaps 0;  
 Oy 507 GAGCCCGCATCTCTGGGAGCCCGGAGACCGGAGAGAGTGGCCGGAGGAAAGC 566  
 Db 4149 GAGCCAGTACATCAATGATTCATGACGACATCTGCGGCCATGAGAGAGAGAGCT 4208  
 Oy 567 GGTCCCGCAGCCGAGAGGAGGAGGAGGATTTGAGCTGAGGAGGAGGAGGAGGAG 626  
 Db 4209 GGTCTAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4268  
 Oy 627 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 686  
 Db 4269 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4328  
 Oy 687 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 746  
 Db 4329 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4388  
 Oy 747 AGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 794  
 Db 4389 CGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4436  
 RESULT 10  
 AAQ84658  
 ID AAQ84658 standard; DNA; 7175 BP.  
 XX  
 AC AAQ84658;  
 XX  
 DT 01-DEC-1995 (first entry)  
 XX  
 DE Human neuronal calcium channel subunit alpha 1B-2.  
 XX  
 KW Calcium channel subunit; antagonist; agonist; diagnosis;  
 KW Lambert Eaton Syndrome; ss.  
 XX  
 OS Homo sapiens.  
 XX

|                           |   |                                  |
|---------------------------|---|----------------------------------|
| FH                        | Key   | Location/Qualifiers              |
| FT                        | CDS   | 144..6857                        |
| FT                        |   | /*tag= a                         |
| FT                        | misc_feature  | 6633..7175                       |
| FT                        |   | /*tag= b                         |
| XX                        |   | /note- "identical to alpha 1B-1" |
| PN                        | M09504822-A.  |                                  |
| PD                        | 16-FEB-1995.  |                                  |
| XX                        |   |                                  |
| PF                        | 11-AUG-1994;  | 94MO-USO9230.                    |
| XX                        |   |                                  |
| PR                        | 11-AUG-1993;  | 93US-0105536.                    |
| PR                        | 05-NOV-1993;  | 93US-0149097.                    |
| XX                        |   |                                  |
| PA                        | (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.                          |                                  |
| XX                        |   |                                  |
| PI                        | Ellis SB, Gillespie A, Harbold MM, Mccue AF, Williams ME,           |                                  |
| XX                        |   |                                  |
| DR                        | WPI; 1995-090900/12.  |                                  |
| DR                        | P-PSDB; AAR71006.   |                                  |
| XX                        |   |                                  |
| PT                        | DNA encoding human calcium channel sub-unit(s) - used for           |                                  |
| PT                        | developing prods. for studying calcium channels, e.g. for           |                                  |
| PT                        | obtaining agonists and antagonists                                  |                                  |
| XX                        |   |                                  |
| PS                        | Disclosure; Page 149-160; 285pp; English.                           |                                  |
| XX                        |   |                                  |
| CC                        | DNA encoding the alpha 1B subunit was isolated by screening a       |                                  |
| CC                        | human basal ganglia cDNA library with fragments of the rabbit       |                                  |
| CC                        | skeletal muscle calcium channel alpha 1 subunit-encoding cDNA.      |                                  |
| CC                        | A portion of one of the positive clones was used to screen an       |                                  |
| CC                        | IMR32 cell cDNA library. Clones that hybridised to the basal        |                                  |
| CC                        | ganglia probe were used to further screen an IMR32 cell cDNA        |                                  |
| CC                        | library to identify overlapping clones that in turn were used       |                                  |
| CC                        | to screen a human hippocampus cDNA library. A series of clones      |                                  |
| CC                        | to span nearly the entire length of the nt. sequence encoding       |                                  |
| CC                        | the human alpha 1B subunit was obtd. Nucleic acid amplification     |                                  |
| CC                        | of specific regions of the IMR32 cell alpha 1B mRNA yielded         |                                  |
| CC                        | additional segments of the alpha 1B coding sequence. A full-        |                                  |
| CC                        | length alpha 1B DNA clone was constructed by ligating portions      |                                  |
| CC                        | of the partial cDNA clones. Nucleic acid amplification analysis     |                                  |
| CC                        | of IMR32 cell RNA and genomic DNA using oligo primers corresp. to   |                                  |
| CC                        | sequences located 5' and 3' of the stop codon of the DNA encoding   |                                  |
| CC                        | the alpha 1B subunit revealed an alternatively spliced alpha        |                                  |
| CC                        | 1B-encoding mRNA in IMR32 cells. This second mRNA product is the    |                                  |
| CC                        | result of differential splicing of the alpha 1B subunit transcrip   |                                  |
| CC                        | t to include another exon that is not present in the mRNA corresp.  |                                  |
| CC                        | to the other 3' alpha 1B cDNA sequence that was initially isolated. |                                  |
| CC                        | The alpha 1B subunit encoded by a DNA sequence contg. an additional |                                  |
| CC                        | exon is referred to as alpha 1B-1 and given in AA084657/R1005,      |                                  |
| CC                        | whereas the other form is referred to as alpha 1B-2 and is given in |                                  |
| CC                        | AA084656/R1006. Following the sequence of the additional exon in    |                                  |
| CC                        | alpha 1B-1 the alpha 1B-1 and alpha 1B-2 sequences are identical.   |                                  |
| XX                        |   |                                  |
| SQ                        | Sequence 7175 BP; 1415 A; 2197 C; 2168 G; 1395 T; 0 other:          |                                  |
| Query Match               | 3.8%; Score 49; DB 16; Length 7175;                                 |                                  |
| Best Local Similarity     | 48.0%; Pred. No. 0.11;  |                                  |
| Matches 201; Conservative | 0; Mismatches 215; Indels 3; Gaps                                   | 2                                |
| QY                        | 250 GAGCGTGCAGGTGAGGTGAACCGCGGCCGGAGAGCCTTCGATGCAATTTCATCGGGCGC     | 309                              |
| DB                        |   |                                  |
|                           | 2751 GAGGCCCCCGAAGCGCGAGAGCGGGGAGCCCGGTGTCGCCGAGGAGGCGCGCGGCAC      | 2810                             |
| QY                        | 310 GAGCAGCTGCAGCTCGAGGGTGCCGAGCCAGCCCGAGAGCCCAGATCGGTTCTCTTCAA     | 369                              |
| DB                        |   |                                  |
|                           | 2811 GCGA-GCCACAGCAAGAAGAGCCGCGGGAGCC--CCCGAGAGCGCGGAGCGAGCCGCGCGA  | 2867                             |
| QY                        | 370 CCCCGTGGCCACGCGCGCGCGGGAATCCCGCGCATCTCGGAGAACCGAGACCCTGTC       | 429                              |
| DB                        |   |                                  |
|                           | 2868 GGCGCCAGGGCCGCGCGCGCGCACCGCGCGCGCGTCTCCCGAGAGAGCGCGCC          | 2927                             |





CC to identify overlapping clones that in turn were used to screen a  
CC human hippocampus cDNA library. In this way, a sufficient series of  
CC clones to span nearly the entire length of the nucleotide sequence  
CC encoding the human alpha 1B subunit was obtained. PCR amplification  
CC of specific regions of the IMR32 cell alpha 1B mRNA yielded  
CC additional segments of the alpha 1B coding sequence. A full-length  
CC alpha 1B cDNA clone was constructed by ligating portions of the  
CC partial cDNA clones (see AA037817, AA037818). Alpha 1B-1 and alpha  
CC 1B-2 are derived by alternative splicing of the alpha 1B subunit  
CC transcript.

SO Sequence 7362 BP; 1446 A; 2276 C; 2214 G; 1426 T; 0 other;

Query Match 3.8%; Score 49; DB 14; Length 7362;  
Best Local Similarity 48.0%; Pred. No. 0.12; Indels 3; Gaps 2;

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGGTGAGTGAACCCGCGCGGACGCTCGTGCAGTGTTCACCTCGGCGC 309  
DB 2751 GAGGCCCCGAGAGCGGAGCGGAGCGCGGTCGCCCGGAGAGAGCGCGCGCCGAC 2810  
QY 310 CGCAGCGTGGAGCTGCGAGGTGCCAGCGACCCCGGATCGGGTTCTGTCAA 369  
DB 2811 CGCA-GCCACAGCAGAGAGCGCGGCGCC--CCGAGAGCGCGAGCGCGCGCGA 2867  
QY 370 CCCGTTGGCCAGCGCGCGCGGAGATCCCGGATCTGCGACAGCCGTAGCCCGTTC 429  
DB 2868 GCGCCAGCG 2927  
QY 430 TCGTCCGTGACCTTGTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489  
DB 2928 GAGCGGAGCG 2987  
QY 490 CCCAGCAGAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549  
DB 2988 GCGCGCGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047  
QY 550 GTGGCGCGCAGAGAACCGGTCTCCCGACCGCGAGAGCGAGCGCGATGTTCTGTCA 609  
DB 3048 GCGGAGAGCG 3107  
QY 610 GCGCAGCGCGCGGTGGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 668  
DB 3108 CACGAGGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 3166

RESUIT 15

AA084657  
ID AA084657 standard; DNA: 7362 BP.

AC AA084657;

DT 01-DEC-1995 (first entry)

DE Human neuronal calcium channel subunit alpha 1B-1.

KW Calcium channel subunit; antagonist; agonist; diagnosis;

KW Lambert Eaton Syndrome; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 144..7163

FT m1sc\_difference 6633..6819

FT m1sc\_feature 6820..7362

FT note= "Identical to alpha 1B-2"

PN MO9504822-A.

PD 16-FEB-1995.  
XX  
XX 11-AUG-1994; 94MO-US09230.  
PF  
XX  
XX 11-AUG-1993; 93US-010536.  
PR 03-NOV-1993; 93US-0149097.  
XX  
XX (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.

PI Ellis SB, Gillespie A, Harpold MM, Mccue AF, Williams ME;  
XX  
XX WPI; 1995-090900/12.  
DR P-PSDB; AAR71005.

PT DNA encoding human calcium channel sub-unit(s) - used for  
PT developing prods. for studying calcium channels, e.g. for  
PT obtaining agonists and antagonists

PS Disclosure: Page 138-149; 285pp; English.

CC DNA encoding the alpha 1B subunit was isolated by screening a  
CC human basal ganglia cDNA library with fragments of the rabbit  
CC skeletal muscle calcium channel alpha 1 subunit-encoding cDNA.  
CC A portion of one of the positive clones was used to screen an  
CC IMR32 cell cDNA library. Clones that hybridised to the basal  
CC ganglia probe were used to further screen an IMR32 cell cDNA  
CC library to identify overlapping clones that in turn were used  
CC to screen a human hippocampus cDNA library. A series of clones  
CC to span nearly the entire length of the nt. sequence encoding  
CC the human alpha 1B subunit was obtained. Nucleic acid amplification  
CC of specific regions of the IMR32 cell alpha 1B mRNA yielded  
CC additional segments of the alpha 1B coding sequence. A full-  
CC length alpha 1B cDNA clone was constructed by ligating portions  
CC of the partial cDNA clones. Nucleic acid amplification analysis  
CC of IMR32 cell RNA and genomic DNA using oligo primers corresp. to  
CC sequences located 5' and 3' of the stop codon of the DNA encoding  
CC the alpha 1B subunit revealed an alternatively spliced alpha  
CC 1B-encoding mRNA in IMR32 cells. This second mRNA product is the  
CC result of differential splicing of the alpha 1B subunit transcript  
CC to include another exon that is not present in the mRNA corresp.  
CC to the other 3' alpha 1B cDNA sequence that was initially isolated.  
CC The alpha 1B subunit encoded by a DNA sequence cont. an additional  
CC exon is referred to as alpha 1B-1 and given in AA084657/R71005,  
CC whereas the other form is referred to as alpha 1B-2 and is given in  
CC AA084658/R71006. Following the sequence of the additional exon in  
CC alpha 1B-1 the alpha 1B-1 and alpha 1B-2 sequences are identical.

SO Sequence 7362 BP; 1445 A; 2278 C; 2214 G; 1425 T; 0 other;

Query Match 3.8%; Score 49; DB 16; Length 7362;  
Best Local Similarity 48.0%; Pred. No. 0.12;

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGGTGAGTGAACCCGCGCGGACGCTCGTGCAGTGTTCACCTCGGCGC 309  
DB 2751 GAGGCCCCGAGAGCGGAGCGGAGCGCGGTCGCCCGGAGAGAGCGCGCGCCGAC 2810  
QY 310 CGCAGCGTGGAGCTGCGAGGTGCCAGCGACCCCGGATCGGGTTCTGTCAA 369  
DB 2811 CGCA-GCCACAGCAGAGAGAGCGCGGCGCC--CCGAGAGCGCGAGCGCGCGCGA 2867  
QY 370 CCCGTTGGCCAGCGCGCGCGGAGATCCCGGATCTGCGACAGCCGTAGCCCGCTTC 429  
DB 2868 GCGCCAGCG 2927  
QY 430 TCGTCCGTGACCTTGTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489  
DB 2928 GAGCGGAGCG 2987  
QY 490 CCCAGCAGAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549  
DB 2988 GCGCGCGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047

Mon Jan 20 09:05:59 2003

us-09-844-864-1.rng

Page 12

OY 550 GTGCGCCGCGAGAAAGCGGTCCTCCCGACCCGCGAAGGAGAGAGGGGATGTCAGGCTCA 609  
 Db 3048 GCGGAGACCGGGGAGAGACCCGCGCGCGGACACCGGGCCGGCACAAAGCCAGAGCTACT 3107  
 OY 610 GGGCAGAGCCGGGTGGAGACAGACCAACACCGGAGAGACGGAAACAGTGGCGCCGAT 668  
 Db 3108 CACGGAGGTGTGGAGAAAGACACACGAGAGAAAGAGAGCCACGAGAGAGAGGCTGAGAT 3166

Search completed: January 19, 2003, 10:25:45  
Job time : 508 secs